

Surrogacy falsely in the dock

A single case of surrogate motherhood seems about to stampede the British government into hasty legislation on human embryology. It should think again.

EVERYBODY is rightly concerned at the wider application of new techniques in human embryology. In most places, there is a need for legislation of some kind. In Britain, for more than the past two years, the government has been planning with creditable care to do something about the problem. Early in 1983, it appointed the Warnock Committee to suggest what might be done, and had a set of detailed if arguable recommendations (see *Nature* 312,389; 1984). But although the committee's chairman, Dame Mary Warnock, has now been appointed to the House of Lords, nothing much else has happened. The government is supposed to have been busy consulting interested parties, but the beginning of this parliamentary session brought no promise that legislation would surface before the summer recess (if ever).

The danger in this delay has always been that the government's capacity to act deliberately would be preempted by the parliamentary device of a private member's bill, the procedure by means of which members of the British House of Commons can advertise their attachment to peripheral causes (the abolition of fox-hunting), canvass opinion on others (capital punishment) and, sometimes, secure legislation on matters so laden with political hazard that governments fight shy of them (whence British legislation on abortion and, last year, the censorship of home video-films). Already such a bill has been introduced to outlaw the scientific investigation of human embryos, although it is too soon to know whether that will muster the support needed to survive. But now, mystifyingly, the government seems to be encouraging just such a piecemeal approach to legislation dealing with just one aspect of the Warnock proposals, the recommendation that surrogate motherhood should be forbidden by law.

The circumstances are curious. Since before the publication of the Warnock report, a private agency in the south of England has been advertising its readiness to provide a surrogacy service on commercial terms, offering to find (and reward) women willing to carry children for childless couples after insemination with sperm from the presumably fertile male partner. The first such case has received a good deal of publicity in the past few months, with a pregnant woman explaining to newspaper and television reporters why she had taken part in the procedure. But when her child was born last weekend, mild public curiosity gave way to zealous public indignation. The local authority in the London suburb where the maternity hospital is situated applied for and won temporary care of the infant (and may be given permanent rights at a court hearing due this Friday), the London police announced that they were investigating whether the birth had been attended by "illegal circumstances", various public dignitaries announced that surrogacy is "just like prostitution", the mother allowed herself to be swept off by a popular newspaper to which she is said to have "sold her story" and the Minister of Health, Mr Kenneth Clarke, promised a "fair wind" for any private member's bill that would ban surrogate motherhood. Then he thought again, and said he would take action off his own bat. Mr Clarke, like most of those concerned, should think yet again.

Surrogacy (Warnock's term), in the sense that led to last weekend's birth, is neither novel nor, in principle, different from the widely practised and condoned technique of artificial insemination with the use of donor sperm (AID). There is plenty of anecdotal evidence of women who grow children for their infecund sisters. As after AID, the child of such a gestation is only

half the genetic product of the parents who claim it as their own, and thus is technically illegitimate. Ironically, the Warnock Committee recommended that AID children should automatically be legitimized as the offspring of their half-unnatural parents while (with the dissent of a tiny minority) inconsistently recommending that surrogacy should be made a criminal offence. But why should a remedy for childlessness which is open to couples in which the male is infertile be denied to those in which the female is infertile? Mr Clarke had better watch out, or he will have the whole women's movement on his doorstep.

Old-fashioned surrogacy does of course entail more taxing problems than AID. What happens, for example, if a surrogate mother wishes to break her contract? Nine months is a long time, while human gestation is physiologically if for no other reason an emotional business. The majority of the Warnock Committee took the view that the emotional risks to surrogate mothers would be greater than the social benefits to others, but was also offended (as British public dignitaries have been these past few days) that money would change hands. The plain truth, last summer as now, is that it makes no sense to ban practices that make people feel uncomfortable by laws that cannot be enforced (as Britain has discovered to be the case for prostitution, homosexuality and abortion). The only prudent course to follow with surrogacy is to ensure that the practice is regulated properly. And the same principle should apply to the other items on the Warnock agenda.

On surrogacy, the steps that need to be taken are obvious enough. First, agencies (which may sometimes be individual physicians) offering to provide these services should be appropriately registered, partly for legal reasons (to establish parentage, for example) and partly so as to ensure that agreements between a surrogate mother and her child's potential parents are equitable. (Everybody worries about what should be done if the mother changes her mind, but what if the prospective parents choose to pull out, perhaps because of congenital malformation?) Far from requiring that money should not change hands, prudent regulation would ensure that potential surrogate mothers and their dependants are financially protected from the still not negligible risks of childbirth. Dark hints from the London police of illegality were in part a reference to the sensible requirement of British adoption law that children for adoption cannot be put up for sale, a difficulty that could be avoided by giving properly registered surrogacy contracts a legal status from the outset (from which it follows that surrogacy agencies should be registered as if they were adoption agencies).

The danger in the situation that has developed is that Mr Clarke will be diverted by the present fuss from paying proper attention to the urgent need for action on the other questions Warnock has dumped on his desk. Surrogacy of the second kind, an embryo formed from gametes of the intended parents gestated in a rented womb, should depend on medical evidence of necessity (as with abortion). Legitimizing AID children is a crying need, especially as many are now falsely registered at birth, but there is a greater need than Warnock allowed that donors should be registered (and that genetic data about them should be retrievable if need be). The same principle should apply to *in vitro* fertilization. The British government's guiding principle, well illustrated by last weekend's fuss and what may yet flow from it, is that the general opinion jumps most readily to prohibition when it is taken by surprise. □



British space agency?

If the British research councils can no longer afford space research, who will pick up the tab?

FOR an eminent astronomer to suggest that Britain should have its own space agency smacks of naivety when sterling is weaker than ever before and when the British research councils are thinking increasingly unthinkably about the harsh decisions that confront them. It may also seem like special pleading on behalf of one section of the community. Budgetary pressures have concentrated the collective mind of the British science community, but divisively. Materials scientists, for example, who rightly complain of insufficient support, wrongly point the finger at "big science". They should know better than to abet the destruction of colleagues in other fields who are just as hard-pressed (see p.91).

A new space agency is not so inopportune a proposal as may first be thought. A panel under the chairmanship of Professor Mark Richmond recently spent several months pondering the costs and the benefits for science, industry and training that accrue from Britain's modest involvement in space (see *Nature* 312, 92; 1984). Its principal recommendations, predictably tame, were that two committees should be set up, one within the Science and Engineering Research Council (SERC) and one at a national level, to coordinate and stimulate potentially interested parties. Another recommendation was that participation in the European Space Agency (ESA) should take priority over other civil space activities. As it happens, the panel was set up by SERC and could not expect directly to influence government ministers. Even so, given the positive achievements and the potential benefits that the panel catalogued, its report was insufficiently robust. Committees such as those proposed are easily ignored. But given money to spend and direct responsibility to a government minister, they have a chance of being audible, visible and effective, which has been shown by the Alvey organization in information technology.

The proposal for a British agency for space is being urged by Professor Martin Rees* and would help do much of what the Richmond panel wanted. At the least, it would exploit the European Space Agency (ESA) for Britain's best national advantage and, at the most, perhaps in a more prosperous future, serve Britain's industrial, research and other communities more directly.

Can such a proposal make sense? In the present financial climate, probably not, but only because the present financial climate itself makes insufficient sense. SERC's dwindling resources cannot support both the core sciences and the increasing demands of lively and intellectually first-rate international collaborations. For what it is worth, the Department of Trade and Industry spent three times as much as SERC in ESA in 1983, primarily on communications satellites and remote sensing. But, by the nature of ESA's constitution, the department was able to do this only because SERC forked out the mandatory science budget required to keep British membership alive. A single channel for British contributions to ESA would avoid this nonsense; obviously Rees considers the risk that science would be swamped by technology to be one worth taking. But what is at stake is only small beer. Britain's total spending on civil activity in space in 1983 was about £80 million, compared with £200 million from West Germany, £300 million from France and £4,500 million from the United States.

What will happen to Rees's proposal is anybody's guess. Probably nothing, at least if past form is any guide. But the British government has to decide something; the question of participation in the US space station programme must be settled at next month's meeting called by ESA in Rome. The government should ask itself two questions. Does Britain have a technological future? And will space at some time be an important arena for civilian activity? If the answer to both questions is "yes" then, as

Rees appropriately asks, can Britain afford to wait for 50 years before deciding what to do?

Management is best

The British government's wish to get rid of poor teachers is not a substitute for good management.

SIR Keith Joseph, said to be an abstemious man, is a glutton for punishment. Even before the full consequences are known of his defeat on his proposal to increase the cost of higher education to middle-class parents so as to find extra for British research (see *Nature* 6 December, p. 483), he was telling school-teachers (last week, at the annual northern education conference) that he will press ahead with a scheme for assessing the performance of working teachers, for rewarding those judged to teach well and for "weeding out" those whose performance conflicts with the educational interests of their students. The minister's difficulty, on this occasion as on others, is that he has identified a serious problem, or part of it, but failed to find a workable solution.

Talk of the assessment of teachers' performance has been in the air for more than a year, since Sir Keith published a white paper on the subject. That teachers probably differ in their competence is not in itself surprising. Reasonable people would be astonished if all teachers were equally able, or if there were not among them some whose performance is an actual impediment to the education of the young. This must be especially likely in a school system, such as the British, in which opportunities for in-service training are far too few, and in which teachers are not required to take advantage of them. But reasonable expectation is flatly contradicted by the assumptions on which British schools are run, and in particular by the peculiarly British convention that a professionally qualified teacher is king in his or her classroom. With the gradual transfer of responsibility for schools from central government to local authorities in the wake of the 1944 Education Act, teachers have won the right to defend themselves against charges of incompetence by simple reference to their professional qualifications. For practical purposes, they have life-long tenure. Further to undermine incentives to performance, both teachers and head teachers are paid on nationally negotiated salary scales with only minor and diminishing rewards for merit.

Sir Keith Joseph's proposed attacks on this cosy system has understandably raised a howl of protest. His case is all the more difficult to sustain because he has not explained how performance will be assessed. There are obvious dangers, not least that teachers will be assessed not by their own performance but by that of their students in public examination—one of the few numerical measures available. The mistake is the assumption that there can be a simple yardstick of performance, or that people's values in professional posts is susceptible to abstract assessment independently of the circumstances in which they work. It would be far better—but perhaps more difficult—that the minister should give head teachers the right and responsibility of deciding how much which teachers contribute to the attainment of a school's objectives coupled with a commensurate say in how individuals should be rewarded. Those who teach at management schools at British universities will readily advise the minister that he cannot hope to build an efficient teaching force without management structure of this kind.

The trouble is that the same principle should apply not only to the schools but to higher education. Academics' first reactions will be to say that the objective measurement of academic performance is impossible, given the diversity of the work academics do. And that view is correct. Some academics will also protest that attempts at the assessment of performance are an interference with academic freedom, which is less cogent. There is no reason why universities should not have a clear idea of their objectives, and civilized ways of ensuring that people contribute to those ends. Such devices are indeed essential if universities are to become more diverse in the years ahead. Nationally negotiated age-related scales of pay have become a hindrance. □

**Britain's Future in Space*, Argo Venture, 18 Victoria Park Square, London E2 9PF, £1.00.

Ten-metre telescope

Caltech wins funds for largest instrument

Los Angeles

CALIFORNIAN astronomers last year went a-courting for funds to build the world's largest optical telescope and last week announced they have found a sweetheart of a deal.

In what is being called the biggest private grant ever made to a scientific project, the William M. Keck Foundation of Los Angeles has pledged \$70 million to the California Institute of Technology to build a 10-metre telescope on Mauna Kea in Hawaii. The University of California will join the project by providing \$2 million for operating costs once the telescope is completed in the early 1990s.

It is time to build a new generation of ground-based optical telescopes, said Dr Marvin L. Goldberger, president of Caltech. The five-metre Hale Telescope on Mount Palomar (operated by Caltech) "has reached its theoretical limits", he said. New technologies such as charge-coupled devices permit the Palomar instrument to collect up to 80 photons per 100. But the maximum limit of detection efficiencies has been reached and astronomers need larger telescopes to gather more light.

The Keck Ten Meter Telescope will be four times more powerful than the Hale Telescope and will be able to detect light from objects that are 12,000 million years old, compared with Palomar's limit of 8,500 million years old.

The new instrument will use new technology to reduce cost, size and weight. Instead of a single polished mirror, the Keck Telescope will consist of 36 hexagonal mirrors six feet wide and three inches thick.

Each mirror will be made using a new technique called "stressed mirror polishing". A mirror blank is warped to a predetermined configuration by applied force and is then polished to a spherical surface.

After polishing, the forde is released and a mirror elastically assumes the desired non-symmetric surface. Dr Jerry Nelson, the astronomer at the University of California who developed the technique said the mirror surfaces resemble "potato chips".

The mirrors are then combined into a single light-collecting device by a computer-controlled aiming system. "We are replacing mechanical rigidity with electronic stiffening", Dr Nelson said. The computer measures the position of each mirror 300 times each second and automatically corrects the position of the array, even in high winds. The new observatory will be located at 13,600 feet on a ridge on Mauna Kea close to a number of other astronomical installations including a British telescope.

The Keck Ten Meter Telescope is probably only the first to win funds in a line-up of other large ground-based telescopes now proposed. The University of Texas has hopes of building a 7-metre instrument and the National Science Foundation has long discussed the idea of building a 15-metre telescope using four mirrors.

The Space Telescope, to be launched in 1986, should be complementary to the Keck instrument, which will have 17 times more light gathering capacity "and will be used to study in detail objects that might have been spotted by the Space Telescope", Dr Goldberger said.

The William M. Keck Foundation was formed in 1954 by the founder of Superior

Oil Company. Howard B. Keck, a son, is chairman and president of the foundation, which now has assets in excess of \$500 million. The foundation decided to give \$70 million to Caltech, foundation directors said, in a continuing spirit of helping scientific endeavours. It has, for example, given \$7 million over a period of years to the Colorado School of Mines.

The Keck Telescope was to have been called the Hoffman Telescope after Marion Hoffman, widow of a wealthy automobile importer Maximilian Hoffman, who had pledged \$36 million to the University of California to build the Ten Meter Telescope. But after Mrs Hoffman died last year, said UC president David Gardner, Caltech was approached to see if both universities could work together to raise the \$85 million needed to complete the project. The Hoffman family will decide before 1 March on continued participation, but there are some who fear that Mrs Hoffman's offer will be withdrawn.

Sandra Blakeslee

Biotechnology regulation

White House holds the ring

Washington

THE White House Office of Science and Technology Policy (OSTP) says, in a major statement on biotechnology regulation*, that the Recombinant DNA Advisory Committee (RAC) of the National Institutes of Health should continue to be the main overseeing body for biomedical research, and endorses the actions of the Environmental Protection Agency in moving to regulate industrial biotechnology.

The question uppermost in the minds of industrial biotechnology companies has been how the Environmental Protection Agency (EPA) would plan to use the Toxic Substance Control Act (TSCA). This act is a catch-all giving EPA the authority to control substances other than pesticides and drugs not covered by other legislation; novel microorganisms are counted as new chemical substances subject to the act's provisions.

There was some concern last year when preliminary drafts circulated by EPA indicated that TSCA could be applied to intermediates used in the manufacture of food products or drugs, which would then be subject to a 90-day notification requirement. EPA now says that such intermediates will not be reviewed under TSCA, leaving the Food and Drug Administration and the Department of Agriculture to deal with such problems as product contamination. In addition, small quantities of chemicals produced exclusively for research purposes will not be subject to the 90-day notification requirement.

EPA has still not resolved, however, some important aspects of when a micro organism constitutes a new chemical substance. Operational definitions are used as far as possible, the guiding principle being

that a new substance is produced if the manufacturing process brings together traits not normally found together in nature. But there are obvious difficulties: for example, EPA seeks comments on whether undirected mutagenesis plus artificial selection should be regarded as giving rise to a new substance. Products or organisms isolated through artificial selection alone, and plants and animals other than microorganisms, are not classed as new substances. EPA has not committed itself on notification requirements for small-scale field testing of novel microorganisms other than pesticides; an interim policy on testing microbial pesticides was announced in October (see *Nature* 25 October 1984, p.695).

Although satisfied that no new legislation will be needed to control the new biotechnology in the foreseeable future, OSTP does see a need for better expert advice in several federal agencies having relevant authority: "the current scientific review apparatus is not designed to respond to all the scientific issues surrounding commercialization of biotechnology". In addition to RAC, OSTP proposes four new independent expert committees, one each in EPA, the Food and Drug Administration, the Department of Agriculture and the National Science Foundation (NSF). The NSF committee would concentrate on the environmental effects of developments in biotechnology; the others would address mainly commercial applications. The work of the committees would be coordinated by a Biotechnology Science Board reporting to the assistant secretary for health.

Comments on the OSTP proposals are welcomed, and should be received before 1 April.

Tim Beardsley

**Federal Register* 49, 50878-50907 (1984).

Nuclear weapons tests

Australian inquiry in London

THE London hearing of the Australian Royal Commission investigating the safety of the British nuclear tests in Australia in the 1950s opened last Thursday with a strong complaint that the British government had failed to live up to its promise of cooperation with the commission's work. Mr Justice James McClelland said that the conduct of the government did not match its assurances given in advance.

The British government decided only in December to be legally represented before the Australian commission. The decision seems to have been made reluctantly, after the commission publicly accused the British government of dragging its feet.

Mr Justice McClelland said that "since the British know so much more than we do about what they were about in our country at that time", cooperation should mean not simply giving access to a mountain of documents, but "giving positive assistance in bringing to light anything of relevance which those documents may disclose". McClelland added that evidence to be heard suggests that the British government told the Australian authorities virtually nothing about what they were doing in Australia during the tests. He also expressed surprise that the British government had insisted that the Australian government must waive its right to prosecute witnesses for perjury, as a condition of allowing the commission to sit in Britain. Such limitations, he said, "sit uneasily with the offer of cooperation and disclosure already made". Mr Robin Auld, for the British government, guaranteed that relevant documents would be produced, and that witnesses including key British personnel responsible for the tests would be made available.

Six ex-servicemen gave evidence to the commission on Monday. Mr Barrie Roberts, of Walsall, had been posted to the Maralinga range, where he said he did not recall being given any particular warnings about the radiation hazard. "We were issued with a handbook which contained all sorts of information, including a section on health. It told you how to deal with snake bites, and warned you about drinking spirits in desert temperatures, but there was nothing in the health section about radiation."

Mr Norman O'Fee, of Eynesbury (Cambridgeshire), joined the Royal Air Force in 1947, and in 1955 was with 24 Squadron flying RAF Hastings aircraft in support of tests at Maralinga. After one of the "Buffalo" explosions, the first air-drop, he flew a group of scientists over ground-zero for about two hours. The plane carried test equipment, and criss-crossed the area at various altitudes. "We were wearing ordinary flying clothing with no safety gear. I could see quite clearly the circle of glazed sand beneath us", Mr O'Fee told the hearing, saying that he had developed

cataracts in both eyes between 1981 and 1983. He considers these may have been due to radiation.

Mr Michael Stephens, one of a meteorological team of eight, joined HMS *Campania* in 1952 and sailed to the Monte Bello islands for what was later confirmed by Dr William Penney (now Lord Penney), one of the scientists on board, to have been an atom-bomb test. Mr Stephens described on Monday a lecture he attended at

which Dr Penney explained the need for the tests, saying that "I remember nothing specific about the danger from radioactivity; nothing he said alerted me to any possible hazard from radiation".

The hearing was told on Monday that a tentative list of about 26 witnesses for the British government will be available to the commission. The British government was also formally asked to produce logs of the Varsity aircraft used in the Buffalo and Antler test series. The Varsityes carried out flights intended to monitor the test sites for aborigines some time before detonations took place.

Susan Watts

South Pacific tests

Mitterrand in the dingo-house

WHILE Britain faces Australian ire over the effects of nuclear testing in Australia in the 1950s (see p.86), France is also in the dingo-house over its continued testing of nuclear weapons on Mururoa atoll near Tahiti. Last month, the temperature of the Franco-Australian conflict was raised when President François Mitterrand accused Australia of solving the problem of the aborigines "by killing them".

This seems to have been a tit-for-tat remark after Australian criticism of French "colonialism" in New Caledonia (where many French Algerians emigrated after Algerian independence, and where some Caledonian aborigines were recently ambushed and shot). But it seems certain to backfire, increasing Australian pressure on France to stop underground testing of 8-10 nuclear weapons a year.

So far, Australia has applied only one sanction to France: the postponement of shipments of uranium, amounting to a few hundred tons of ore a year. But the sanction is weak. France has recently diversified its sources of supply and, moreover, there is a glut of uranium on the market. In fact, Electricité de France, the French utility buying the Australian uranium, readily agreed to the postponement, which will last until France stops bomb-testing in the Pacific.

Although further sanctions are not for the time being contemplated in Australia, 18 Pacific nations including Australia and New Zealand are drawing up an international convention which should make the joint economic zones of signatories a "nuclear free area", to exclude nuclear dumping, storage and weapons testing. The terms of this convention are not likely to be softened by M. Mitterrand's accusations.

Nevertheless, when tempers cool, some thought might be given to recent scientific assessments of the environmental impact of the Mururoa tests by a scientific mission from New Zealand, Australia and Papua New Guinea, which allayed immediate concern about short-term effects. Moreover, an exhaustive recent review* concluded that:

● For the South Pacific region, the average annual effective dose equivalent from natural sources of ionizing radiation is approximately 2,000 microsievert, only half the world average;

● Exposure to artificial sources, mainly the radionuclides formed during weapons tests in the atmosphere, is on the average perhaps two to three times lower in the South Pacific than for the world as a whole.

The exceptions are Niue Island, "a documented example of unusually high natural radioactivity", Guam (which may also have high natural levels) and certain of the Marshall Islands still contaminated by fall-out from US nuclear tests.

However, while the UNEP report considers the French underground tests to be safer than tests above ground, "one should be particularly concerned at the possible long-term effects, such as leakage of radionuclides into the ocean, especially if the testing programme... is to continue".

The UNEP technical group preparing the report, chaired by Dr Michael Bacon of the Woods Hole Oceanographic Institution in the United States, believes that past environmental assessments and publication of results on Mururoa previous to the recent scientific mission have been "inadequate". Bacon considers it "interesting" that the French recently began conducting tests within the Mururoa Lagoon, rather than in the atoll itself, as was past practice. French scientists are concerned with the danger of tsunami-like waves produced by sediment slides on the atoll slopes.

Nevertheless, the UNEP technical group concludes that "the present nuclear weapons testing... involve[s] only a small, quite possibly non-existent risk to human health and the environment in the South Pacific region". The group says that there is little "scientific basis for judging these activities to be unacceptable", but acknowledges that "legal, political and moral principles" may be more relevant.

Robert Walgate

*Radioactivity in the South Pacific United Nations Environment Programme Reports and Studies no. 40 (1984).

French research

All civil servants now

FRENCH researchers working for the major national research councils are about to receive the pieces of paper that declare them to be civil servants and so guaranteed a job for life, nearly a year after the principle was first agreed by the French National Assembly. The ministry of public administration, it seems, has been fighting a rearguard action to prevent researchers being given too many exceptions to civil service rules, for fear of other civil servants demanding the same conditions, but now finally the battle is over, and last week the research councils published details of the new arrangements.

These new contracts of employment consist "mostly of exceptions", according to the Centre National de la Recherche Scientifique, the major council for basic science. Yet the contracts are "on the whole not as dynamic as we'd have liked" according to Claude Kordon, a researcher with the Institut National de la Santé et de la Recherche Médicale, who has pursued the reform of scientific employment ever since the present socialist government came to power in 1981.

According to Kordon, one of the main disappointments is with the categorization of technicians, who will be confined to 13 separate categories and career patterns, according to their qualifications. French laboratory technicians are demoralized by the lack of opportunity for initiative and advancement, Kordon believes, and he had been campaigning for merely one or two categories to allow for more movement and promotion. But the ministry had been "completely obstinate".

Among the exceptions, however, are some "positive things" such as a degree of scientific assessment of technicians for promotion, and certain opportunities to shift categories. But these might affect only a small proportion of jobs, as the whole French research system—like that in many other countries—is blocked by a mass of staff recruited in the 1960s and not due for retirement for at least a decade.

As for researchers themselves, they will now be divided into two classes (*chargés* and *directeurs* de la recherche) as opposed to the present three (*attachés*, *chargés* and *directeurs*), to match a similar change in the categorization of university researchers. Previously, all researchers were on rolling three (or more) year contracts, which could in principle be revoked (for example if the Centre National de la Recherche Scientifique pulled out of a line of research) but in practice never were (except for gross misconduct). Now the jobs are officially for life. Even if the French research councils were abolished tomorrow, the state would have to pay its researchers and technicians their salaries and pensions to their deaths.

To the security-conscious French, this could be a passport to greater mobility,

Kordon believes—and mobility is certainly something French research lacks. It is also encouraged by elements in the new contracts that offer bonuses for researchers who change geographical location or subject of research.

In the final analysis, however, the new contracts may prove to be little more than paper, leaving the *status quo* intact until the great retirement of the 1990s. More important may be the introduction of a

three-year PhD degree, which will begin in earnest this October, and supplant the present one-year diploma of advanced study (DEA), "third-cycle" two-year thesis and a "state thesis" often not completed before the age of 30. The existing system frequently tied researchers down to working in the same laboratory for upwards of a decade in what might be the most productive years of their lives. A three-year PhD, followed by a search for jobs, might loosen the French system considerably. But not, of course, until 1988, when the first French PhDs will begin to emerge.

Robert Walgate

AIDS tests alarm blood banks

Washington

THE American Association of Blood Banks (AABB), while acknowledging great uncertainties in the tests now being developed to screen blood for AIDS (acquired immune deficiency syndrome) antibody, has recommended that all blood-banks use the tests on every unit of blood collected and that donors whose blood tests positive be informed of the results.

A number of blood bank officials have expressed serious reservations about using the tests because of the high false-positive rates reported in initial trial runs (see *Nature* 13 December 1984, p. 583). In addition, little is known of the clinical significance of a positive result; it is possible that many with antibody to the AIDS virus, HTLV-III, may never contract the disease. The AABB recommendation notes that "only after several years of experience,

along with appropriate epidemiological studies, will there be sufficient information to assess the meaning of a positive test for antibodies to HTLV-III".

According to AABB, because of the false-positive problem, the Food and Drug Administration, which must approve any test before it is put to routine use, is planning to require that some kind of confirmatory test be run on all positive samples. The most likely such test would be a Western blot, according to AABB and blood-bank officials.

FDA approval of the first test kits is expected shortly.

Although the AABB recommendation is not binding on its membership (which includes 2,400 hospital and community blood banks in the United States), its standards are widely followed throughout the world.

Stephen Budiansky

North Korea

Home science seems best

NORTH Korea must bring its science and technology up to world level, says a leading political monthly, *Kulloja*. In particular, technical education must be improved, including "teaching the science and technology of other countries in accordance with the specific situation" in North Korea.

This last proviso has apparently caused certain problems. Although the article is signed by a certain Pak Yong Ghol, its somewhat confused and inconsistent argument suggests joint authorship and/or subsequent ideological revision. Several paragraphs extol the burgeoning of science elsewhere. The article nevertheless concludes that such achievements must be studied "only for the purpose of enabling us to know what we have that is better and in order more properly to carry out revolution and construction".

In effect, the only practical goals of the proposed drive are fairly mundane and include a non-specific exhortation that science and technology must be used to solve the "problems concerning the Chucheization of the country", which is a reference to reconstruction according to the Chuche variant of socialism favoured by

the Party leader, Kim Il-sung.

The tenor of the article, that foreign science and technology, however advanced, "cannot readily fit the situation in our country" appears, at first glance, simply a repetition of the truism that for a developing country, middle technology is more appropriate than advanced. (So why devote so much space to the sophisticated developments which find practical application elsewhere?) In the context of North Korea, the concept seems to have considerable political overtones. Koreans, it is implied, can study best in Korea.

Last month, this proposition was expounded in the somewhat unlikely setting of an international conference of journalists in Pyongyang, where a contrast was drawn between the South Korean attitude, which favours studying abroad, and that of North Korean Party Secretary Kim Chong-il, who refused the chance to study at the "best university" in an unspecified "socialist country" (almost certainly the Soviet Union), preferring to study "on the basis of the reality of the fatherland in which the Chuche idea was being embodied".

Vera Rich

US research ships

Naval build-up in prospect

Washington

REPLACEMENT of the ageing fleet of US research ships has received a surprise commitment from the US Navy, but there are still uncertainties about the continuation of state-of-the-art research into the 1990s.

Almost all the 20 ocean-going ships operated by universities and by research institutions such as Woods Hole and Scripps were built with federal funds; most are at least 15 years old, and some are as much as 28 years old. The average lifetime of a research vessel is considered to be 25 to 30 years. Moreover, Bob Dinsmore of Woods Hole Oceanographic Institution, who is heading a committee studying fleet replacement for academic users, says that some comparatively young ships "are already approaching obsolescence" in their suitability for modern research.

To the surprise of many, the Navy recently seemed to reverse a 20-year trend of diminishing interest in the academic research fleet. A policy statement signed by the Secretary of the Navy, John Lehman, sets out a number of new initiatives in oceanography, including a commitment to seek support in fiscal year 1987 for the construction of one new research ship for academic users. In addition, the statement signalled an intention to develop a long-range plan to replace the other four large ships originally built by the Navy.

But researchers remain unassured. Professor James McCarthy, an oceanographer at Harvard University, says "the Navy is very hard to predict". Many things could happen between now and 1987; and even if the \$35 million that the Navy is now talking about for construction does survive in the Defense Department's budget request a year from now, it would still have to be approved by the Office of Management and Budget, the President and the Congress.

The National Science Foundation (NSF), which has shown a steadier commitment to the academic research fleet, has however found itself up against financial constraints that have ruled out any thoughts of building a new large ship. Only one of the six large research ships in the academic fleet was built by NSF (*Atlantis II*, operated by Woods Hole); and while NSF has been able to underwrite construction of five new smaller coastal ships since 1970, an increasing share of its oceanography budget has gone on operating and maintaining the existing fleet. (Overall, NSF supports some 75 per cent of federally supported oceanography research, the Navy about 15 per cent and the National Oceanic and Atmospheric Administration, the National Aeronautics and Space Administration and the US Geological Survey the balance.) Maintenance costs have escalated over the past decade as more and more replacement parts for the older ships have had to be

specially made; fuel costs have taken a large bite as well. Sandra Toye, the NSF official who deals with oceanographic facilities, says the agency has no plans for constructing new ships in the next few years.

From the researchers' point of view, the problem is both the age of the ships and the increasing sophistication of experiments, which require larger scientific crews, more on-board instrumentation and larger laboratory space. "By the 1990s, we will be desperately short of ocean-going vessels", McCarthy says, unless construction of replacement ships begins at once. "We're currently using vessels that are on their second lives." Four of the increasingly-in-demand large ships were built in the early 1960s, the other two, *Knorr* (operated by Woods Hole) and *Melville* (operated by Scripps), in 1970.

The trend in recent years has been towards greater use of interdisciplinary research teams, of perhaps 25-30 people, studying such questions as atmosphere-ocean interactions. Satellite navigation and communications equipment is being used more; real-time data analysis, which allows the researchers to modify their experiments during a cruise, is also becoming the norm.

The Committee on Fleet Replacement for the University/National Oceanographic Laboratory System (UNOLS), which Dinsmore is heading, has put together a "wish list" that may carry considerable weight in Navy and NSF planning. Similar calls for replacement of the fleet are coming in a series of National Academy of Sciences reports due out this year. The UNOLS committee is proposing that the Navy begin construction of three large general-purpose ships in 1987, 1988 and 1989; these would replace the three oldest Navy-built ships, which date from the early 1960s. In addition, the Navy would be asked to refurbish and modernize the two ships it built in 1970, extending their life to 2010. Another five ships would come from NSF: one general-purpose ship to be built in the second half of the next decade, two special-purpose geophysics vessels, one submarine support ship and one polar ship—an item long sought by the research community.

The Navy and NSF have together supported conceptual design studies for the new ships; model-testing of two designs is planned for this winter, and UNOLS expects to select candidates for a detailed preliminary design in March. The Navy has expressed interest in the the so-called SWATH design—a two-hulled ship that would be very stable.

A selling point for the US replacement effort that is sure to be heard in the coming year is that Britain, France, West Germany, and the Soviet Union have all constructed well-equipped, state-of-the-art research vessels within the past few years.

Stephen Budiansky

Bacterial boycott

Strains withheld from Soviets

Washington

BIOLOGISTS in the United States and Britain are urging their colleagues in the West not to supply their Soviet counterparts with bacterial strains for research until refusnik scientist David Goldfarb is allowed to leave the Soviet Union. Goldfarb had been assured last April that he would be permitted to emigrate, but after resigning his position and making final preparations to depart, he was abruptly informed by KGB agents that his permission to emigrate had been suspended and that he was under investigation for attempting to smuggle out of the country "secret" bacterial strains.

Almost all of Goldfarb's collection of auxotrophic mutants of *Escherichia coli* K-12 were originally obtained from US and British researchers. Western scientists familiar with Goldfarb's research say it is "very innocuous" and totally unrelated to military applications. Goldfarb's visa had been arranged by the Soviet Academy of



Sciences, which apparently by-passed the KGB; the incident may be part of a larger tug-of-war between Soviet biologists and the KGB over KGB efforts to impose secrecy upon recombinant DNA research.

The attempt to suspend shipments of strains to the Soviet Union is being organized by Max Gottesman of the National Institutes of Health and Charles Yanofsky of Stanford University in the United States, and Michael Yudkin of the University of Oxford and Simon Baumberg of the University of Leeds in Europe. Gottesman was unable to say how many requests for samples US and European laboratories receive from Soviet researchers; he said his own laboratory had received a few requests a year from Soviet scientists until recently.

Goldfarb, who is 66, has one son living in Israel and one in the United States, who now teaches at Columbia University. That son, Alexander, who has been in contact with Goldfarb, is supporting the boycott, believing that it may succeed in putting pressure on the Soviet authorities.

Stephen Budiansky

US nuclear power

Diablo Canyon at full power

Washington

A *cause célèbre* of the US nuclear power industry appears finally to have been laid to rest—or almost. On the last day of 1984, a US Court of Appeals panel gave its opinion that the Nuclear Regulatory Commission (NRC) acted within its legal discretion in granting a full power operating licence to the Pacific Gas and Electric Company for its Diablo Canyon 1 plant, ordered in 1966. With two minor exceptions, which did not merit judicial relief, the panel dismissed a petition by Mothers for Peace in San Luis Obispo alleging that the commission had committed legal errors, and proclaimed that “Diablo Canyon is ready to begin generating electric power for the citizens of California”. The decision is being hailed as a significant victory by the industry.

In fact, the plant has been producing electric power for the citizens of California since November, though at only 50 per cent of capacity; Pacific Gas hopes to move to full power operation by the end of March. But San Luis Obispo Mothers for Peace have not yet given up: they plan to use their legal right to take the petition to a full hearing before all 12 justices of the court. And there is always the Supreme Court.

The issue of an operating licence for Diablo Canyon became controversial in the 1970s, with the discovery of an offshore geological fault within 3 miles of the site. An element of farce was added in 1981 when it was discovered that blueprints used in construction of the reactor had mistakenly been reversed; an investigation then revealed numerous other design faults. A full power licence was eventually issued by NRC last August, only to be suspended immediately by the Appeals Court after the petition from Mothers for Peace.

The Court of Appeals panel now finds that NRC has been guilty of only two minor misdemeanours, relating to operator training on simulators and the right of petitioners to a hearing on construction quality assurance. Both issues were later overtaken by events and have no present bearing on safe operation of the reactor, according to the panel. On the question whether earthquakes near the plant pose a safety threat, which has been central to the Diablo Canyon saga, the panel states that the seismic design of the plant is such that “the likelihood that an earthquake will trigger a nuclear accident at the facility is so small as to be rated zero”. The chance of a core meltdown occurring coincidentally with an earthquake—relevant to emergency evacuation plans—is “extremely small”, and NRC acted within its discretion in deciding that this eventuality need not be considered in evacuation planning.

Evacuation plans were, however, the subject of a dissenting opinion by Judge Patricia Wald, who held that ignoring the

effects of earthquakes for emergency planning was “a substantial lapse of rational decision making”. Judge Wald confessed to being “baffled” by NRC’s “paradoxical stance” on the importance of earthquakes, considered in plant design though not in evacuation planning.

Nuclear industry spokesmen are, not naturally, hoping that the Diablo Canyon ruling will signal a turn-around in the industry’s fortunes. A recent comparison study by the Atomic Industrial Forum (AIF) indicates that new nuclear power plants ordered in the late 1980s will typically have a 10 per cent price advantage per kilowatt hour over coal-fired plants. The study, which used assumptions conservative for the nuclear case, highlights the long construction times for nuclear plants in the United States—up to 16 years for more recent plants. In France, by comparison, the range is 5 to 8 years. The study group concludes, however, that standardization of licensing procedures, improved

economic conditions and the recent rapid growth in electricity demand—up 6 per cent in 1984 over the previous year—will all reduce US construction times to the best European levels. Demand is now more than 2 per cent above the 1984 level, and surplus capacity is down to 30–32 per cent. Demand for new plants will rise sharply when surplus capacity falls to 20–25 per cent, according to AIF; some surplus is needed to cope with periods of peak demand.

The AIF conclusions depend on nuclear fuel maintaining its current cost advantage against coal. But the indications are that the fuel price advantage will tend even more strongly towards nuclear electricity.

Apart from the Diablo Canyon ruling, 1984 as a whole was a mixed year for the industry. NRC imposed 27 fines on electricity utilities for safety violations. And the total of eight low-power operating licences issued during the year exactly offset the number of cancellations. But that is fewer cancellations than in some recent years, and the industry—now providing 13 per cent of US electricity—expects to bring 30 more plants into operation before the end of 1986.

Tim Beardsley

Bhopal disaster

Technical inquiry under way

Bhopal

THE 16.2 tonnes of methyl isocyanate (MIC) left over at the Union Carbide India Limited (UCIL) plant after last month’s tragedy, when more than 2,500 people died, has been converted into the pesticide Carbaryl and the plant has been declared safe. UCIL staff, released temporarily from police custody, disposed of the remaining MIC after a specially assembled team had been satisfied that the operation would pose no risk. A third of the people living in Bhopal nevertheless moved out of the city during the five-day operation (codenamed “Faith”), which went off uneventfully. Meanwhile, the state government of Madhya Pradesh made it clear that the plant would not be allowed to reopen either in Bhopal or elsewhere in the state.

The cause of the accident on 2 December is the subject of a judicial inquiry and another by the Central Bureau of Investigations (CBI). Dr S. Varadarajan, director of the Council of Scientific and Industrial Research and former head of Indian Petrochemicals Limited in Baroda, is helping with the CBI investigation. He says that, at the time of the accident, the temperature of the MIC storage tank had exceeded 80 degrees centigrade and that the pressure rose above the 2.5 kg per square centimetre that the relief valve could deal with.

The cause is thought to have been either the entry of water into the tank or the spontaneous polymerization (in the absence of inhibitors) of the liquid MIC, which had been in storage for a month, a longer period than normal. The heat released during the process of polymerization might

be sufficient to vapourize the remainder of the MIC. A maintenance team had been cleaning the vent lines with water shortly before the accident, but it is not known how two tonnes of water (the minimum that would have been needed to trigger such a reaction in the 45 tonnes of MIC in the tank) could have entered through the apparently airtight seals of the tank.

A third theory is that nitrogen used to pressurize the tank might have been contaminated; MIC is known to react with virtually all chemicals. The only evidence collected so far consists of samples of polymerized MIC recovered from the tank by forcing through high-pressure nitrogen.

Between 80,000 and 90,000 people are undergoing treatment for MIC exposure. About a third of the victims have returned to hospitals with secondary complaints of liver and kidney disorders. The problems appear to be similar to those caused by chlorine poisoning, in which case, according to Dr Hans Weill, professor of pulmonary medicine at Tulane Medical School, New Orleans, “a large number will recover”. But an Indian expert on chest diseases, Dr A.S. Paintal, says that the victims would be prone in future to all kinds of allergic disorders.

Although some people have been blinded, most of the eye injuries are thought to be self-healing because only the outer layer of the cornea (the epithelium) would be involved. But Indian eye surgeons nevertheless estimate that 1,000 donor eyes will be needed for corneal grafts in those whose cornea has become opaque.

K.S. Jayaraman

Japanese telecommunications

Now for the sale of NTT

Tokyo

WITH the Christmas-time passage of three telecommunications bills through the Diet, Japan is entering 1985 ready to follow the United States and Britain and end the state monopoly of telecommunications services. Later this year, Nippon Telegraph and Telephone (NTT), a state corporation with an annual income of 4.5 million million yen (US \$19,000 million) and 320,000 employees, should be up for sale in what is almost certain to be the largest offer ever made on any stock market, capping even that of British Telecom in November 1984.

Not surprisingly, the passage of the bills has put the Japanese telecommunications scene into a state of extraordinary excitement. Big companies are hurriedly trying to find groupings that will give them enough muscle to enter the telecommunications market as serious competitors to the denationalized NTT. Foreign telecommunications companies, including the US giants AT&T and IBM, are preparing for their first chance to enter the value-added network market on equal terms, together with US satellite makers that will soon be able to offer telecommunication satellites to private companies.

The three bills that turn NTT into a private corporation also differentiate between those telecommunications businesses that actually own and operate telecommunications services (common carriers) and those that lease circuits from them to offer other kinds of services. To ensure that the actual means of transmission remains in Japanese hands, foreign companies are to be excluded from the first category. Just which Japanese companies will be big enough to get into it is the subject of intense speculation.

Obviously, the privatized NTT will be the leader in the field; indeed, the government will have to restrain its activities in order to stop it crushing new competitors. Four other groups have emerged as contenders for a "second NTT". First and foremost is a group headed by Kyocera, an up-and-coming high-tech ceramics company. Sony, Mitsubishi and more than 200 other companies are ready to team up with Kyocera. Next come the national railways and the national highways public corporations. What they both have to offer is routes — either railways or freeways — right into the centre of big cities along which optical fibre can be laid cheaply. Finally, there is a group organized by Keidanren (the Federation of Economic Organizations) which intends to purchase the nation's first private telecommunications satellite and to have it in operation by 1988.

Among second category businesses — those offering services on lines they have leased — the Japanese scene will be transformed by the entry of foreign companies.

Almost every major US company doing value-added network (VAN) business (providing the computers and software that will allow many different kinds of computer, desktops to mainframes, to talk to one another through telecommunications circuits) has found a Japanese partner to help sell its services. The only remaining worry (for the United States) is that the telecommunication bills retain licensing powers for large VAN operators and that there could be discrimination in favour of home companies.

The legislation allowing free competition comes into effect on 1 April. When and how much of NTT will be sold off is not yet clear, however. Autumn looks a likely time and eventually the government is committed to retaining only one-third of the

stock. The Ministry of Posts and Telecommunications, made a last-ditch attempt to keep a large part of the proceeds from the sale for itself with a plan to set up no less than four new telecommunication research and development agencies. The plan was stamped on by the Diet, however, which intends to put the proceeds towards paying off government debts. There, alas, even the vast sums released by the sale of NTT will be just a drop in the ocean.

The total market value of NTT shares could exceed Y10 million (\$41,000 million) but the Japanese deficit has grown to a staggering Y110 million million — 40 per cent of GNP. Each year almost 20 per cent of the government's total revenue goes towards paying interest on money it has already borrowed. The sale of NTT might perhaps just make it possible for the government to stop borrowing money to pay interest on money it has already borrowed.

Alun Anderson

AMPTE

Christmas fireworks a success

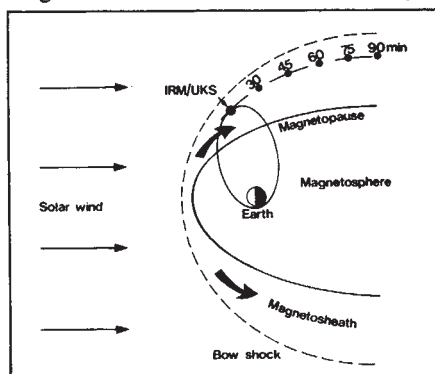
THE artificial barium comet released on 27 December (two days later than first planned) as part of the AMPTE mission (Active Magnetospheric Particle Tracer Explorers) seems set to yield much of the information expected of it. Although the conditions for the release were not as expected, two of the AMPTE spacecraft were able to measure the effects both inside and outside the cometary cloud. Moreover there is enough barium left to run the experiment again under different conditions.

The cloud was released by the West German Ion Release Module (IRM). The 1.25 kg of barium, which was quickly photo-ionized, expanded to a diameter of about 250 km, while the tail extended to 11,600 km. At the time of release, a separate spacecraft, the UK Subsattellite (UKS), was about 150 km from IRM. The spacecraft were situated in the magnetized plasma that envelopes the Earth, originating from the Sun and interacting with the geomagnetic field. Once ionized, it was expected that the cloud, being highly conducting, would "expel" the ambient magnetic field. The interest lies not only in

this process but also in the way the external magnetic field re-enters the dissipating cavity thus formed, as well the electromagnetic process occurring near the cloud boundary and the overall behaviour of the ion cloud.

Preliminary indications are that the experiment was a success. Both spacecraft observed intensification of the ambient magnetic field by up to a factor of six or more, and IRM observed a drop in the interior field down to the noise level. The speed of the solar plasma was about 2 million km per hour, but this was halved near the cloud. A large variety of plasma oscillations was observed and the energy of electrons in the solar wind rose from 10 eV to 1 KeV. The solar wind flowed around the cloud, drawing particles away to form the tail, while the pressure on the front caused the cloud to accelerate downstream.

The positions of both spacecraft relative to the comet were not as originally planned. At their height (100,000 km above the Pacific) and position at the dawn flank of the Sun-Earth reference frame, they were expected to be within the magnetosheath (see figure). This is the region of plasma flow immediately downstream of the "bow shock" where the supersonic solar wind is first affected by the geomagnetic field. But the experiment has highlighted the extremely variable position of the bow shock. At the time of the experiment it was much closer to the Earth than anticipated and both spacecraft became enveloped in the supersonic solar wind, to the detriment of the UKS's proximity to the cometary tail. However, both spacecraft measured significant effects; furthermore the flow conditions are "cleaner" in the solar wind than in the magnetosheath, and the repeat experiment (perhaps in June) should be all the more valuable as a result. Philip Campbell



The comet release mission as originally planned (see text). The positions of the cloud as it is pushed downstream to the nightside are marked.

Environmental fall-out

THE details are likely to be settled later in the month of a radical plan for the reorganization of the British Natural Environment Research Council (NERC), the smallest of the four largely scientific research councils in Britain, with a budget exceeding £90 million a year. One of the objectives is to divert a larger proportion of the annual spending on research towards the universities. Another is to make the research programme based on the council's own laboratories more flexible by concentrating the direction of the programme at the council's Swindon headquarters.

NERC, which has been under financial pressure for several years and has suffered particularly from the steady decline of commissioned research from government departments, has been mulling over its new plan for close on a year. The point has now been reached when trades unions and institute directors are being consulted about

the prospects for the next five years, during which it is thought that the research council might shed up to a third of its employees, not all by retirement and resignation.

The impending change is not universally welcomed. Although some academic researchers in fields such as geophysics consider that the new strategy will be advantageous, many of those employed at council institutes insist that research direction at a distance (from Swindon) will not bring the economies and flexibility expected. The position of the British Geological Survey seems to be especially precarious, given the degree to which its work is supported by government departments.

NERC itself insists that details of the plan are not yet complete, and that they cannot be settled until discussions with the unions have been completed. At least some council members are not at this stage committed to all the details of the plan. □

Venture capital

British industry expands

HOPES that British economic fortunes will revive with the greater availability of venture capital were illuminated but not settled last week, when Morgan Grenfell Ltd, the London merchant bank, announced that it had successfully raised £30 million to launch a new venture capital fund, Advent Capital Ltd. Two years ago, two other venture capitals funds (called Advent Technology and Advent Eurofund) were launched with the encouragement of Monsanto, the international chemical company. There seems no doubt that British investors in venture capital funds will benefit from such investments. To judge from the prospectus issued with last week's placing, it remains in doubt whether investment in British companies is the best course.

Stock market rules require that closely linked companies should not replicate themselves without describing how the original has fared financially. According to last week's prospectus, Advent Technology now has a capital worth of £15.3 million (against £10 million in March 1981) and Advent Europe a value of £12.7 million (compared with £10 million in May 1982).

The investment portfolio of Advent Technology, for example, shows that the substantial shifts in the value of investments (greater than one third) were in British Agricultural Genetics Ltd (which recently went public), Felton Fluid Ltd (a British company now worth half the original cost), KWE Inc (a US microwave component manufacturer, now worth four times its original value). These and other smallish investments almost cancel out, but

the great appreciation in the value of Advent Technology (and to a lesser extent, of Advent Europe) arises from a profit of nearly £3 million from the sale of the British company Xenotron Ltd. □

AFRC secretary

Dr John Jinks, professor of genetics at the University of Birmingham, has been appointed secretary of the Agricultural and Food Research Council from 1 May 1985 in succession to Sir Ralph Riley, who retires



at the end of April. He will join an embattled research council, which has for the past five years been under pressure to manage its affairs on less, and which is at present engaged on a "restructuring" programme that will entail the loss of at least two of its research institutes. □

UK nuclear physics

Research board chairman quits

ONE of the first personal casualties of the shortage of research support in Britain was the resignation earlier this week of Professor D.C. Colley, a member of the Science and Engineering Research Council (SERC) and chairman of the council's Nuclear Physics Board, which is among other things responsible for the cost of the British subscription to the European Organization for Nuclear Research (CERN), near Geneva.

In a strongly worded letter to Sir Keith Joseph, Secretary of State for Education and Science, Colley (who is professor of high-energy physics at the University of Birmingham) says that he is resigning from the council in protest at the way in which funds have been allocated to nuclear physics by the Advisory Board for the Research Councils (ABRC), the committee on whose advice the government shares out the annual research budget.

Colley particularly objects to the way in which the research councils are not to be compensated, in the year ahead, for the increased cost of international subscriptions at a time when sterling is falling rapidly, saying that ABRC has "surrendered an important point of principle" in not pursuing an agreement reached last year with the British Treasury. Colley says that the annual dispute over the cost of international subscriptions is divisive as between different fields of science, and that Britain, so far as he knows, is the only European country in which extra funds to pay for exchange rate fluctuations must be met from within the overall science budget.

Colley also complains that support for nuclear physics research has been further constrained by the instruction, issued by ABRC, that none of the extra money recently made available to SERC should be spent on nuclear physics. In his letter, Colley says the result will be a further decrease of £1 million in funds for nuclear physics. Colley goes on to say that he can only interpret ABRC's decisions as "part of a deliberate attempt" to pre-empt the recommendations of the Kendrew committee (due to report in two months on continued British membership of CERN), saying that it "ill befits" the largest research council to act in a way that prejudices the outcome of an inquiry it has itself set up.

Of SERC policy as a whole, Colley says that he is increasingly concerned about the policy of diverting funds away from "big science", apparently in the belief that it is "SERC's job to save the United Kingdom economically". He argues that funds amounting to only 5 per cent of the British research and development budget can have "at most a marginal impact" and that the council of SERC has been "grossly shortsighted" in following this course. □

THE Nature biotechnology stock index is being reconstructed. Monthly publication will be resumed as soon as possible.

Is science value-free?

SIR — Contrary to the view of Bernard D. Davis (*Nature* 27 September, p. 294), value judgements are inherent in scientific knowledge and scientific practice. They can be identified most readily when science is applied to problems of the real world such as energy, environment and health. These problems are very complex and scientific information on them is generally far from complete. As a consequence, the scientist inevitably has to make value judgements which involve the selection of data, of scientific procedures for analysing data and of assumptions about unknown variables. Furthermore, in presenting the results of this selection process to peers, to decision-makers or to the public, the scientist has to make selections of definitions and terminology and of the context in which the results are exhibited. (Davis's use of the value-laden phrases "periphery of science", "fashionable", "confusion" and "radical left" should be noted.)

Such value judgements do not occur in random directions. They are part of a science which is performed in a particular social, political and economic context, which includes:

- industrial, bureaucratic, government and military motivation for, and financial support of, most scientific research;
- the limited accessibility, understandability and exploitability of most scientific research to anyone except large organizations with vested interests;
- the effective monopolization of the opportunity to do scientific research by full-time professionals;
- the restriction of decision-making about scientific research priorities to a small group of power-holders.

In this context, the value judgements, and therefore science itself, are given a direction or bias, namely that which is selectively useful to elite groups. (See Hilary and Steven Rose (eds) *The Political Economy of Science* and *The Radicalisation of Science*; Brian Martin, *The Bias of Science*; Randall Albury, *The Politics of Objectivity*.) In many examples — for example the risk of fallout from nuclear weapons tests, electricity generation planning — the public is given the incorrect impression that the issues involved are purely "scientific", "technological" or "economic" in nature, when in reality they rest in part on social, political or ethical judgements (see my article in David Oldroyd (ed.) *Science and Ethics*, NSW University Press).

Why is the notion that science is "objective" or "value-free" promoted so strongly? I would suggest that an important reason is that science is used widely to legitimize the products and policies of elite groups. My own particular bias is towards a kind of democracy which allows greater public participation in decision-making on

social, political and economic issues. By questioning the notion of value-free science, it is my hope that the community will eventually create a new type of science which is more responsive to the needs of the community at large than to the military, big business, elite scientists, bureaucracy and government.

The views expressed in this letter are not necessarily those of any organization with which I am associated.

MARK DIESENDOERF

*CSIRO Division of Mathematics & Statistics,
Canberra ACT 2601,
Australia*

Misnomer

SIR — The article on the prehistory of Amazonian Indians by Robert May (*Nature* 1 November, p.19) included an inference that bronchial pneumonia (*sic*) decimated groups of indigenes.

Bronchopneumonia is a disease affecting the terminal airways, giving rise to patchy consolidation of the lung parenchyma. Lobar pneumonia causes consolidation of whole lobes of the lung. Either type may be spread by infection and cause epidemics.

"Bronchial pneumonia" is a popular misnomer with no pathological basis due, I suspect, to mishearing bronchopneumonia. In this case, I suggest the epidemic was due to influenzal bronchopneumonia as in the pandemic following World War I.

LAURENCE J.R. BROWN

*Department of Pathology,
University of Leeds,
Leeds LS2 9JT, UK*

SAAO telescopes

SIR — If 156 UK astronomers believe that "continued funding of the South African Astronomical Observatory (SAAO) amounts to acquiescence in apartheid" (*Nature* 312, 295; 1984), it is probable that many of these UK astronomers will not use SAAO.

The scientific returns on the investment by the Science and Engineering Research Council (SERC) in SAAO probably correlates with the fraction of active UK astronomers who are prepared to use the SAAO telescopes.

The declaration presented to SERC indicates that a significant fraction of active UK astronomers are unlikely to consider using the SAAO telescopes. That surely is a good enough "scientific consideration" for SERC to transfer its meagre astronomy resources to telescopes that all astronomers are prepared to use.

JOHN BARUCH

*Department of Physics,
University of Leeds,
Leeds LS2 9JT, UK*

Down with animal lib

SIR — *Nature* is increasingly devoting its pages to leading articles, letters, book reviews, news items and lengthy commentaries (for example, refs 1-6) on animal "rights", "welfare" and "abuse". In 30 issues from 26 April to 15 November 1984, about 14 pages were dedicated to this subject compared with fewer than 7 dealing with human rights from Argentina to the Soviet Union. Matters such as the devastating famine in sub-Saharan Africa, to the relief of which scientists can contribute a great deal, were not even mentioned.

I do not care whether cuddly animals (nobody seems to care about prokaryotes, ectothermes and protozoa) are abused, killed, eaten, mistreated or otherwise molested for pleasure, profit, the advancement of science or the American way, within or without laboratories, and I wish *Nature*, if not to adopt my point of view, at least to focus more on the species to which presumably most of its readers and, more importantly, subscribers belong.

DAN GRAUR

*Center of Demographic and Population Genetics,
University of Texas,
Health Science Center,
Houston, Texas 77225, USA*

1. Beardsley T. *Nature* 309, 8 (1984).
2. Cosslett V.E. *Nature* 309, 186 (1984).
3. Goodwin L.G. *Nature* 309, 729 (1984).
4. Stevens C. *Nature* 312, 295-297 (1984).
5. Remfry J. *Nature* 312, 191 (1984).
6. Editorial. *Nature* 312, 2 (1984).

India's diversity

SIR — I take strong exception to your reference to India as a "huge country, which is in reality six or seven countries" (*Nature* 8 November, p. 87), a remark suitable for extremist newspapers not periodicals such as *Nature*. It is difficult to judge if a journal which I so assiduously read was indulging in provocative journalism apparently legitimizing secessionist attitudes or simply exhibiting ignorance. Extending your view, this very island nation would thus represent an artificial conglomeration of essentially disparate peoples of at least four countries or perhaps three discounting a united Ireland. India was strong enough to withstand the loss of Mahatma Gandhi during those critical few years after independence and will weather the recent loss of Indira Gandhi with reinforced national unity amid a proud exhibition of cultural diversity.

M.R. SURESH

*Laboratory of Molecular Biology,
Medical Research Council Centre,
University Medical School,
Hills Road, Cambridge CB2 2QH, UK*

The phrase objected to was not intended in the sense assumed but as a reference to the cultural diversity of India — Editor, *Nature*.

No pattern yet for snowflakes

Attempts to model the growth of highly symmetrical crystal structures are most conspicuous for their failure. The urgent need is for truly cooperative models.

THE problem of calculating the shapes of snowflakes persists. That is the simplest conclusion to be drawn from the best effort so far to apply to the growth of a crystal with intrinsic symmetry the theory of the growth of solid phases by the random aggregation of particles (atoms, say) in circumstances in which growth is limited by the speed of diffusion. For the numerical experiments carried out by Tamas Vicsek from the Institute of Technical Physics at Budapest, Hungary, but working last year at the physics department of Emory University at Atlanta, Georgia, suggest that even in the simplest circumstances (the growth of a two-dimensional crystal on a square lattice), the underlying symmetry is only crudely represented in the shape of the crystal (*Phys. Rev. Lett.* **53**, 2281; 1984). But Vicsek's calculations do at least suggest what might next be tried.

During the past two years, simulations of diffusion-limited aggregation have become extremely fashionable, partly because they are a natural way of generating fractal structures, partly because of the practical importance of the process. The now standard way of carrying them out is conceptually to take a finite piece of some lattice (for example, a piece of a square lattice a few hundred lattice intervals in each direction), to suppose there is a nucleation somewhere in the lattice, to suppose that particles capable of joining a growing aggregate can enter the stage at random points on the periphery and then migrate through the lattice by means of a random walk from one point on the lattice to a nearest neighbour and then, finally, to suppose that such a particle will stay put once it reaches a lattice point adjacent either to the nucleation centre or to a point occupied by the growing aggregate based upon it. The model is simply if tediously calculated (but computer time mounts up to such an extent that nobody seems to have attempted a three-dimensional lattice). The simplest results are those most simply expected. Newly arriving particles stick at the places where they first encounter neighbours. The result is a randomly constructed network of particles which are less densely packed at the periphery than towards the centre, which is another way of describing a structure with fractal properties — the total mass (or number of particles in the aggregate) increases less quickly than the square of the linear dimension of the growing structure. The greater the distance from the nuclea-

tion centre, the fuller the whole structure is of nothing.

The upshot of this kind of stimulation has so far been what might, on reflection, have been expected. The larger the central structure, the much greater the chance that it will intercept the random walk of a newly arriving particle. By the time the aggregation has grown to fill a substantial proportion of the space set aside for it at the outset, the chances that the next random walk will yield a particle that sticks on the side it enters must be high. This is another way of saying that the simulational system is less a good model of diffusion-limited aggregation as the aggregated structures grow. Perhaps practitioners in the field should scale their models to take account of this source of bias.

The problem of geometrical symmetry is much more difficult, if only because it must involve physical considerations. Vicsek seeks to solve the problem by changing the rules of aggregation on a simulated piece of lattice. First, he says, a randomly-walking particle must not stick whenever it comes next to an occupied site, but the probability that it will stick must be a function of some representation of the local curvature, most simply measured (by computers) as the ratio of occupied to unoccupied sites within some given distance. Second, the argument goes, it makes sense to allow newly arriving particles to relax to nearest-neighbour sites if the potential energy will thereby be decreased (or the number of nearest neighbours increased).

The two modifications of the sticking rule that Vicsek introduces are not as simple-minded as they may seem. In the real world of solids and fluids, there is every reason why an aggregating particle's chance of sticking to a growing surface should depend on the most immediately local curvature, for which purpose a simple count of the immediately neighbouring sites which are occupied should be a good approximation to the Gibbs-Thomson macroscopic rule relating aggregation probability to surface tension. That newly aggregated particles should relax to neighbouring lattice sites on a growing aggregate is also commonsensical; at the least, this may be a good simulation of what happens in a reality, the permanent sticking of new particles that reach energetically favoured sites, and the rapid solution (volatilization) of particles that first stick elsewhere.

The snag, in the sequence of four simulations of growth on a two-dimensional

square lattice reproduced with Vicsek's paper, is that the outcome is far from being a structure with fourfold symmetry that would be expected. To be sure, there are signs of predominant growth in each of four axial directions, largely as a consequence of the positive correlation between sticking and curvature, but there is no detailed simulation of the exquisite detailed rotational symmetry embodied in every other ice crystal (where the axis of rotation is, of course sixfold and not fourfold).

Vicsek says in his conclusion that the outcome of his simulations is only qualitative, which is fair enough. Nobody would pretend otherwise of a simulation of a real physical system in which only the interactions between nearest neighbours are counted as significant. The obvious difficulty, of course, is that there can be very little hope of reaching better conclusions by taking account of next-nearest neighbours, or of further refinements of that kind. And a little reflection will show why no other conclusion is possible. The aggregation of particles onto a growing surface will be determined exclusively by local properties, among which surface tension and the opportunities for energetically advantageous migration will be important. But the symmetry of a whole crystal, represented by the exquisite sixfold symmetry of the standard snowflake, must be the consequence of some cooperative phenomenon involving the growing crystal as a whole. What can that be? What can tell one growing face of a crystal (in three dimensions this time) what the shape of the opposite face is like? Only the lattice vibrations which are exquisitely sensitive to the shape of the structure in which they occur (but which are almost incalculable if the shapes are not simply regular).

What Vicsek seems therefore to have accomplished is to have demonstrated the limitations of computer simulation in the representation of high crystal symmetry. The simple sticking rules of the model for diffusion-limited aggregation may serve well enough for dealing with the growth of amorphous particles (soot in the atmosphere, for example) but they are unlikely to throw much light on the reasons why so many simple crystal-growing operations yield such symmetrical structures. Back now in Budapest, Vicsek may feel inclined to think of adding the complexity of lattice vibrations to the problem of simple Ising lattice calculations. But if not him, then somebody. Please.

John Maddox

Plasma physics

Inertial confinement of fusion

from M.H. Key and R.G. Evans.

THERMONUCLEAR fusion offers a potentially limitless source of energy if the extreme conditions required for the fusion reaction can be achieved with cost effectiveness in a fusion reactor. The ignition temperature of the reaction is around 10^8 K and at this temperature the fuel (the hydrogen isotopes deuterium and tritium) is a plasma of free electrons and positive nuclei. The plasma must be confined long enough for the nuclei to collide and fuse, thereby releasing over 10,000 times their initial thermal energy in the form of neutrons, α particles, tritons and protons. Two major approaches have been adopted to achieve sufficiently long confinement. In one, typified by the Joint European Torus (JET) project in Oxfordshire, magnetic confinement is used to try and confine a low density plasma for longer than one second in equilibrium in a magnetically-insulated vacuum vessel. The other, known as inertial confinement (ICF), requires sudden heating of a milligram or so of dense fuel to the fusion temperature with only inertial confinement of the plasma for the $\sim 10^{-12}$ seconds it takes for the explosive expansion. Both approaches were considered in detail at a recent conference* but we shall confine ourselves to the session devoted to laser fusion and ion beam fusion in ICF.

The criterion for sufficiently long confinement, often called Lawson's Criterion, can be expressed as $n\tau \sim 10^{14} \text{ cm}^{-3}\text{s}$ where n is the number of nuclei per unit volume and τ the time of confinement. In the case of ICF, $\tau \approx r/v$, where r is the radius and v the thermal velocity of the ions, and Lawson's Criterion for inertial confinement can be expressed in terms of the density, ρ , as $\rho r \sim 1 \text{ g cm}^{-2}$; also the mass of the plasma scales as ρr^3 . Since Lawson's Criterion gives ρr^3 , it follows that the mass, and thus ignition energy, scale as $1/\rho^2$, high density being required to satisfy Lawson's Criterion for a small fuel mass. The ICF principle is used in the only successful man-made fusion reaction to date – the hydrogen bomb – in which an atomic (fission) bomb provides the thermal energy to compress and heat a large mass of fusion fuel, releasing a huge amount of fusion energy.

The aim of ICF research is to achieve controlled fusion of a relatively tiny mass of fuel at densities as high as 200 g cm^{-3} – one thousand times the density of liquid deuterium. Compression to such high densities requires a minimum pressure of several thousand million atmospheres. It is planned to achieve this in the implosion of

hollow shell targets containing the fuel at liquid density (0.2 g cm^{-3}) and $\leq 10^6$ K driven by a pressure of about 10^8 atmospheres sustained for several nanoseconds. Ignition of the compressed fuel would be achieved at its centre in a small fraction of the mass by a strong spherically-converging shock wave. The whole fuel mass would then 'burn' with reabsorption of the thermonuclear α particles heating the cold fuel. Both lasers and ion beams are being considered as fusion drivers.

Laser fusion has the considerable advantage that it is relatively easy to irradiate small targets at the intensity ($\sim 10^{15} \text{ W cm}^{-2}$) needed to generate a pressure of 10^8 atmospheres. Many aspects of laser fusion have therefore been studied using scaled down targets ($r = 100 \mu\text{m}$), which require only kilojoule laser pulses to drive them. Small-scale experiments continue to be very important in optimizing the compression process, most notably in showing the advantages of short wavelength ($\leq 0.5 \mu\text{m}$) irradiation.

Experimental verification of improved laser absorption and reduced preheating by plasma wave instabilities, for $0.53 \mu\text{m}$, $0.35 \mu\text{m}$ and $0.26 \mu\text{m}$ illumination, was reported from the Lawrence Livermore National Laboratory, the University of Rochester and the Ecole Polytechnique respectively, with numerical simulation data from the US Naval Research Laboratory. In terms of target coupling, it seems that the shorter the laser wavelength the better, and nearly complete absorption is measured at $0.35 \mu\text{m}$ and $0.26 \mu\text{m}$.

To reach the megajoule pulse energy, which calculations indicate is needed to drive reactor-size targets, requires new laser technology. The world's most powerful operational laser at present is in Japan, at Osaka University's Institute for Laser Engineering. The 12-beam glass laser (Gekko XII) generates 50 terawatts in 10^{-10} second pulses at the infrared wavelength $1.05 \mu\text{m}$. By early 1985, US supremacy will probably be restored when the 10-beam glass laser NOVA designed for 100 kJ in 10^{-9} seconds (100 terrawatts) is due to be commissioned at Lawrence Livermore.

Laser fusion experiments with this kind of machine have mainly been of either a 'direct drive' or 'indirect drive' type, as illustrated in the figure. The latter approach has produced a compressed density as high as about 20 g cm^{-3} at Lawrence Livermore but the details have never been published. Therefore there was considerable interest at the meeting in similar work reported by Osaka University group. In their 'cannonball' implosion scheme, the ICF pellet is contained in an outer enclosure, with small holes to allow the input

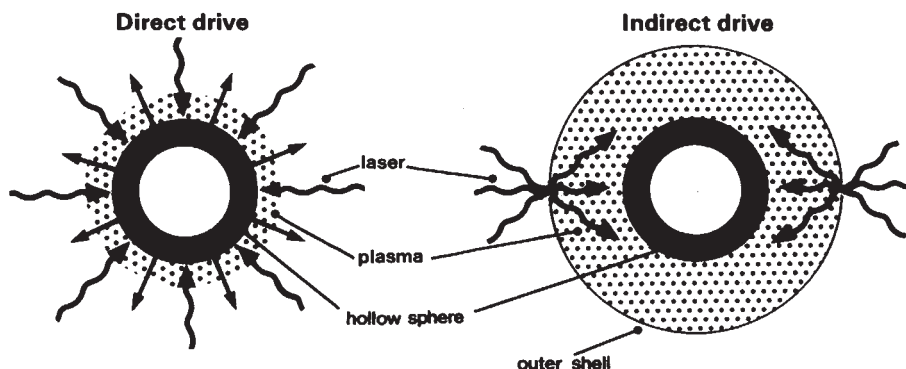
of the laser energy. The inner pellet is driven by the plasma pressure in the space between pellet and outer wall, and/or by X rays generated by the heating effects of laser light at the walls of the enclosure. The Japanese scheme is believed to offer an improvement both in uniformity of pressure and in absorption efficiency but the experiments, which used the relatively long wavelengths of $1 \mu\text{m}$, still showed a significant level of preheat and reached densities of only about 10 g cm^{-3} . A density of up to 40 g cm^{-3} was reported for similar experiments with a wavelength of $0.53 \mu\text{m}$ at the Lawrence Livermore, using the new two-beam prototype of the NOVA laser to generate 10 kJ, 10^{-9} second pulses, but few experimental details were given.

Direct drive compression requires a highly uniform ($\pm 1\%$) irradiation of the target, which is difficult to achieve. The University of Rochester reported high quality laser beam profiles and predicted a $\pm 10\%$ uniformity with their 24-beam OMEGA laser now being converted to operate at $0.35 \mu\text{m}$. The required uniformity of implosion should then be achieved through the 'smoothing' effects of heat transfer within the pellet atmosphere. Direct measurements of this smoothing factor were presented from the Central Laser Facility at the Rutherford Laboratory in the UK and from CEA Limeil, France. Results from the Ecole Polytechnique showed implosions driven by $0.26 \mu\text{m}$ light in small scale experiments; better symmetry was obtained by focussing the laser beams in front of the target, instead of behind as is normal practice. From their attempts to achieve simultaneously the benefits of short wavelength lasers and thermal smoothing, the Rutherford Laboratory presented X-ray radiographic data of very symmetrical implosions driven by either six or twelve beams at $0.53 \mu\text{m}$, in which compression up to 10 g cm^{-3} was achieved.

The optimum wavelength for laser fusion lies somewhere between 0.2 – $0.5 \mu\text{m}$. Shorter wavelengths have the advantage of increased coupling efficiency and reduced preheat, but place more severe constraints on laser uniformity. The new technique of 'induced spatial incoherence', being developed by the Naval Research Laboratory and at Osaka University, may produce more uniform illumination by scrambling the phase of the laser so that, in effect, it becomes a large number of incoherent beamlets. Successful operation of this scheme could make the $0.26 \mu\text{m}$ krypton fluoride laser the hot favourite for an ICF test reactor because of its high efficiency (several percent) and short wavelength. Indeed, a major problem for laser fusion is the low efficiency (0.1%) and low repetition rate (1/hour) on Nd glass lasers compared to the efficiency of about 5% and repetition rate of 50 Hz needed for an ICF power station.

Ion beams are generally more efficient ($\geq 10\%$) and can operate at higher repetition rates. Their exploitation as fusion

*10th international conference on 'Plasma Physics and Controlled Nuclear Fusion Research', Imperial College London, 12–19 September 1984.



Alternative schemes for laser-driven implosions. In direct drive, a hollow spherical target is illuminated with laser light from all directions and plasma flows out from their radiated surface. In indirect drive, the target is contained in an outer shell, with laser radiation entering through two tiny ports. The enclosure is filled with plasma.

drivers has lagged behind that of laser beams because they cannot be scaled down in pulse energy and duration and at the same time drive small targets at relevant intensities. Research has therefore concentrated on building large machines to study the driver technology problems.

The scaling problem, which is less severe when light ions ($Z \leq 10$) are used, was the subject of major presentations from the National Research Laboratory, and Sandia National Laboratory in the US, and from Osaka University. Subtle changes to the anode profile in ion beam diodes have produced power densities up to $10^{13} \text{ W cm}^{-2}$, close to that needed to drive an ICF pellet. These light-ion machines have hundreds of kilojoules of stored electrical energy and use electrical techniques to deliver most of this energy in about 10 nanoseconds. The first stage is a normal pulse-charged transmission line delivering a pulse of about 25 nanoseconds, and a new second stage of pulse compression is now in use in both the United States and Japan. The transmission line is effectively short circuited at the output end by a burst of plasma, resulting in the electrical energy being stored inductively. As the plasma current builds up, it magnetically insulates itself, rapidly shuts off the current and gives an output pulse

twice the input voltage and one third of the pulse duration.

In principle, heavy ions ($Z > 50$) are also a contender in the long-term battle for an ICF driver but the principle still awaits proof on even an intermediate-sized machine of reasonable cost. But the advance of induction linac technology has led to a US proposal for a test accelerator that would come within a factor of three of an ICF driver in all key parameters.

The results presented at the meeting show that magnetic and inertial confinement fusion are now in similar positions. In both schemes the ignition temperature can be achieved at low densities, while high-density relatively low-temperature plasmas are steadily being pushed towards the Lawson Criterion. Perhaps magnetic fusion is closer to demonstrating a controlled fusion burn, but the reactor engineering problems seem less daunting in ICF. To demonstrate an efficient fusion burn, both approaches will probably need bigger and more expensive machines in the next few years. The engineering of a reactor will present major challenges for subsequent decades. □

M.H. Key and R.G. Evans are at the SERC Rutherford Appleton Laboratory, Chilton, Didcot, Oxon. OX11 0QN, UK.

Evolutionary biology

Raising the wrong children

from Paul H. Harvey

ONE of the most important constraints on an animal's reproductive output is its need to invest in its young. Perhaps, then, it is not surprising that patterns of behaviour have evolved by which parents persuade others to raise their offspring. For example, foster-parents may be led unknowingly to invest in offspring that are not their own. Cuckoos and cowbirds are familiar for their use of other species but with the help of some ingenious techniques such parasitism is now being detected within species. Detailed field studies on a number of species of birds and mammals are showing that parentage may be quite

uncertain, even in species where it would seem most secure.

Males and females use different methods to get their offspring accepted by a rival. Males surreptitiously copulate with other males' partners, whereas females dump their young into other females' broods. While the former strategy is found among both birds and mammals, the latter seems largely restricted to birds, possibly because female mammals rapidly learn to recognize their own offspring.

Andersson¹ has identified two evolutionary routes to successful egg dumping. The first occurs when nest sites are few.

Two females lay eggs in the same nest and, when one withdraws to find another nest in which to lay, the other is left with a clutch of mixed maternity but is unable to identify which eggs are her own. She therefore incubates them all. The second route to brood parasitism results from nest failure: if a female loses her clutch to a predator, or if her nest is destroyed before she has completed laying, she may lay in another nest of the same species. Comparative data support Andersson's ideas; nest sites are often limiting among hole-nesting species, a group in which egg dumping is prominent, while heavy rains followed by flooding of nests often leads to increased intraspecific brood parasitism². A recent study by Charles Brown³ identified intraspecific brood parasitism in colonial cliff swallows (*Hirundo pyrrhonota*) and provided an interesting though unexplained correlate with sociality. In colonies with less than ten nests, females laid eggs only in their own nests, but in larger colonies up to 24 per cent of the nests contained at least one illicit egg.

One lesson that emerges from the growing literature on the subject is that extremely detailed observational studies may be necessary to detect egg dumping. In many species, the intruding female removes an egg and replaces it with one of her own so that the returning (now surrogate) parent finds an apparently undisturbed nest. Since most birds lay an egg a day, the casual observer will also record a state of domestic tranquility.

Careful observation of individual birds also shows that the bounds of monogamy or the pair bond are frequently overstepped. But copulation does not ensure paternity and there are still very few paternity-exclusion studies. Some of the most reliable of these use electrophoretically-detectable protein variation to demonstrate that a putative father could not possibly have sired a particular offspring. One of the first such studies on birds has recently been published⁴. It concerns the apparently monogamous eastern bluebird (*Sialia sialis*), and shows that at least 5 per cent of adult males and 15 per cent of adult females in the sample were caring for at least one offspring that was not their own. The same method has been used in three notable paternity-exclusion analyses of mammals. One concludes that 78 per cent of Belding's ground squirrel (*Spermophilus beldingi*) litters are sired by more than one male⁵. The other two studies show that males of two harem-dwelling species — the sparnose bat (*Phyllostomus hastatus*) and the black-tailed prairie dog (*Cynomys ludovicianus*) — could not have fathered all the offspring^{6,7}.

Electrophoresis is not the only way to demonstrate that a male has been cuckolded. Vasectomy is an even more incisive technique. In the classic study of this kind, Bray *et al.* vasectomized male harem-holding red-winged blackbirds (*Agelaius phoeniceus*) and yet found that several of

their mates laid fertile clutches⁸.

A less intrusive technique has recently provided surprisingly good correlative evidence for successful cuckoldry in the pied flycatcher (*Ficedula hypoleuca*)⁹. This species is one of a growing number that is recognized as being polyterritorial. A male defends a small territory around a nest hole, a female mates with him and starts to lay a brood, and the male then dupes a second female into mating with him at a second nesting territory (often some distance from the first) before leaving her, to return and feed his first mate's young. The primary female succeeds in fledging considerably more young than does the second¹⁰. While the male is away from the nest, the primary female sometimes copulates with a secondary male from an adjacent territory. To see if these copulations result in paternity, Alatalo *et al.*⁹ measured tarsus lengths of the primary female, the primary male and the offspring in several hundred pied flycatcher nests. The correlation between mothers and offspring is considerably better than between fathers and offspring. Not surprisingly, perhaps, there is a significant positive correlation between the tarsus lengths of offspring and the adult males in the nearest territory. This correlation is unlikely to be caused by common environmental factors since there is no correlation between the tarsus lengths of adult males from neighbouring territories. Instead, it seems that about a quarter of the hatchlings in the study were not sired by the resident male. A less exhaustive study produced similar results with the congeneric collared flycatcher (*Ficedula albicollis*)⁹.

One clear implication of these studies is that the genetic structure of a population will seldom be quite what one would expect from the idealized picture of its breeding system derived from casual observation. If conspecifics are being fooled, it is likely that human observers are too. The social behaviour of a species, including its breeding system, necessarily evolves in the context of its actual genetic structure. To understand the functional significance of differences in social organization among birds and mammals, we will need yet more detailed accounts of the actual patterns of relatedness in species with different kinds of mating systems and ecologies. □

1. Andersson, M. in *Producers and Scroungers: Strategies of Exploitation and Parasitism* (ed. Barnard C.J.) 195 (Croom Helm, London, 1984).
2. Yom-Tov, Y. *Biol. Rev.* **55**, 93 (1980).
3. Brown, C.R. *Science* **224**, 518 (1984).
4. Gowaty, P.A. & Karlén, A.A. *Behav. Ecol. Sociobiol.* **15**, 91 (1984).
5. Hanken, J. & Sherman, P.W. *Science* **212**, 351 (1981).
6. McCracken, G. & Bradbury, J. *Science* **198**, 303 (1977).
7. Hoogland, J. & Foltz, D.W. *Behav. Ecol. Sociobiol.* **11**, 155 (1982).
8. Bray, O.E., Kennelly, J.J. & Guarino, J.L. *Wilson Bull.* **87**, 187 (1975).
9. Alatalo, R.V., Gustafsson, L. & Lundberg, A. *Oikos* **42**, 41 (1984).
10. Alatalo, R.V., Carlson, A., Lundberg, A. & Ulfstrand, S. *Am. Nat.* **117**, 783 (1981).

Paul H. Harvey is a Visiting Professor in the Department of Biology, Princeton University, Princeton, New Jersey 08544, USA.

Supernova remnants

Defining the Crab nebula

from Virginia Trimble

NO TWO remnants of the stellar explosions we call supernovae are alike. But the Crab nebula (SN 1054) — complete with 33 msec pulsar and lots of helium, yet lacking the usual rapidly-moving, shock-edged supernova remnant shell — is even less alike than most. New observations, presented at a recent symposium*, serve to emphasize its uniqueness. Among the highlights were: dynamical information on the jet that protrudes from the main nebula; the probable demise of earlier evidence for an extended halo; identification of an exceedingly similar, though not identical, remnant in the Large Magellanic Cloud; and discovery of a new class of non-thermal galactic radio and X-ray sources that are neither Crab-like nor shell-type supernova remnants (for more on which see pages 113, 115 and 118 of this issue).

Sydney van den Bergh (*Astrophys. J.* **160**, L27; 1970) first reported a luminous jet extending from the north-east nebular boundary of the Crab nebula. Later work

*The Crab Nebula and Related Remnants, George Mason University, Virginia, 11–12 October 1984.

established that it emits both line (thermal) and continuum (synchrotron) radiation, and prompted countless models, which were invariably a bit contorted as the axis of the jet has never pointed back to the pulsar position, either now or in AD 1054. These models will have to be rethought in light of four new observations. First, the jet is moving outward along its own axis at about 4,000 kilometres per second (R.A. Fesen and T.R. Gull, University of Colorado and Goddard Space Flight Centre). Second, the local magnetic field is aligned with the jet (A.S. Wilson, University of Maryland; T. Velusamy, Tata Institute). Third, while moving outwards, the jet is simultaneously expanding cylindrically and perpendicularly to its axis at 360 kilometres per second (Peter Shull, Oklahoma State University). And, fourth, composition of the jet is quite close to that of the adjacent nebular filaments, including a lower-than-average ratio of helium to hydrogen (R.C.B. Henry, University of Oklahoma). The length divided by the axial velocity and the width divided by the trans-

Milan Hašek 1925–1984

THE death on 14 November 1984 of Milan Hašek has deprived Czechoslovakia of one of its most distinguished scientists and one of the pioneers of transplantation biology. At the time of his death he was a leading scientist in the Czechoslovak Academy of Sciences' Institute of Molecular Genetics.

Hašek's great and enduring achievement was to have discovered the phenomenon of immunological tolerance, independently of all others, by the ingenious expedient of making a vascular bridge — a parabiosis — between two embryonated hens' eggs. The chorio-allantoic membranes of the two eggs were laid bare over an area of about 0.5 cm² and these two windows were apposed to each other using a fragment of embryonic tissue to make a vascular bridge between the two membranes. Hašek showed that the two chickens hatched from these eggs had a greatly diminished ability to make red cell agglutinins against each other's erythrocytes, and later it became clear that the chickens were mutually tolerant of skin grafts that were exchanged between them.

When Hašek began his work, he was a dedicated communist. This was evident from his early papers, in which the results were interpreted in terms of 'negative hybridization', a phenomenon that he believed to cast doubt on 'the absurd theory of the genes'. Hašek was far too intelli-

gent to persist in these doctrinal delusions. Once he became aware of immunogenetic developments in the West — especially the discovery of immunological tolerance in inbred mice — he quickly mastered modern transplantation theory and, with some able colleagues, made many important contributions to it. Indeed, some of his papers have taken their place among the classics of immunology.

Hašek happened to be abroad at the time of the Russian occupation of Czechoslovakia in 1968 but, despite his protest against this tragic event, decided to return. Sadly he was subsequently deprived of his influential post as director of the prestigious and, under his leadership, vigorous Institute for Experimental Biology of the Czech Academy of Sciences. This, together with a recurrent illness, kept him in the wilderness for a long time. In recent years he re-established himself in the Institute of Molecular Genetics and, together with a group of bright collaborators, made numerous and substantial further contributions to transplantation immunology.

Milan Hašek was a giant in more ways than one. He was a man of great physical stature and his enthusiasm and zest for life was legendary and infectious. He will be greatly mourned and missed by friends in the scientific community throughout the world.

P.B. Medawar & L. Brent

	SNR 0540-69.3 + PULSAR	CRAB NEBULA + PULSAR
Pulse period	0.050 s	0.033 s
dP/dt	$4.79 \times 10^{-13} \text{ s s}^{-1}$	$4.23 \times 10^{-13} \text{ s s}^{-1}$
dE/dt	$1.5 \times 10^{38} \text{ erg s}^{-1}$	$4.5 \times 10^{38} \text{ erg s}^{-1}$
Implied surface B	$4 \times 10^{12} \text{ G}$	$3 \times 10^{12} \text{ G}$
I_x (total)	$10^{37} \text{ erg s}^{-1}$	$2 \times 10^{37} \text{ erg s}^{-1}$
Ratio $L_x : dE/dt$	0.05	0.05
X-ray pulsed fraction	23% (increases with E)	4% (increases with E)
X-ray pulse shape	Sinusoidal (split peak)	Sharp pulse, wider interpulse
Age from dP/dt	1,600 yr	1,230 yr
Nebular radius	1 pc	2 pc
Nebular V (expansion)	$1,250 \text{ km s}^{-1}$	$2,000 \text{ km s}^{-1}$
Expansion age	800 yr	844 yr
Real age	800–100 yr (?)	930 yr
Optical pulsed L	$10^{34} \text{ erg s}^{-1}$	$10^{34} \text{ erg s}^{-1}$
Optical pulsed fraction	0.6%	0.4%
$B-V$ colour	0.85 ± 0.35	0.5
Optical pulse shape	Sinusoidal (split peak)	Sharp pulse and interpulse
Radio pulsed flux	$\leq 0.5 \text{ mJy}$	6.6 Jy (10 mJy at LMC distance)
Radio size	$\approx$ Optical remnant	$\approx$ Optical remnant
Remnant composition	O/H high; He/H normal	O/H normal; He/H normal
Estimated progenitor	25–30 $M_{\odot}$	8–10 $M_{\odot}$

verse velocity both give time scales of 600 years. One is left with the impression that something punched a hole at a random weak point on the nebular surface, and that the local mix of thermal and relativistic gas shot out into a less confining medium.

In the past, X-ray, radio, and optical observers have all presented evidence that the Crab might have a normal, but inconspicuous, shell-type supernova remnant around it (see Trimble, V., *Rev. mod. Phys.* **55**, 511; 1983), and Kurt Weiler (National Science Foundation) listed half a dozen such composite remnants. But the shell now seems to be imaginary. The diffuse X rays are due to instrumental and inter-stellar scattering (C. Mauche and P. Gorenstein, Center for Astrophysics, Harvard); even very faint filaments do not show the velocities of more than 2,000 kilometres per second that are needed to reach sizes greater than two parsec in 930 years (R.A. Fesen, University of Colorado and D. A. Ketelson, Kitt Peak National Observatory); and VLA radio maps have not revealed any emission outside the main nebula and jet down to levels about one-thousandth of the central peak surface

brightness (A.S. Wilson; T. Velusamy).

Finally, the pulsar and associated remnant 0540-69.3 in the Large Magellanic Cloud resemble the Crab nebula and its pulsar in many ways. The data, summarized in the table, come from Frank Seward (Center for Astrophysics), Stephen Reynolds (National Radio Astronomy Observatory), John Middleditch (Los Alamos), Robert Kirshner (University of Michigan) and Kenichi Nomoto (University of Tokyo). Some are rather tentative. Significant differences seem to include the X-ray pulse shape and fraction, the nebular composition, and — perhaps — the absence of radio pulses and a dP/dt in 0540-69.3 that is sufficiently erratic to prevent measurement of the slowing-down index, PP/P^2 , which tests pulsar emission mechanisms. Theorists are now hard at work accounting for both the similarities and the differences between the two objects. □

Virginia Trimble is Visiting Professor of Astronomy, University of Maryland, College Park, Maryland 20742 and Professor of Physics, University of California, Irvine, California 92717, USA.

Cell biology

Subtleties of actin assembly

from Albrecht Wegner

ACTIN is practically everywhere. Together with microtubules, actin filaments make up the principal dynamic constituents of the cytoskeleton and they control the shape and movement of cells. Actin is capable of entering into a variety of complex and sensitive interactions, which regulate its state. Even the simplest possible system — that in which the actin monomer (G-actin) interacts with itself to form linear aggregates (F-actin) often several microns in length — has been found to embody such an astonishing number of thermodynamic and kin-

etic subtleties that it has kept many of the sharpest minds and fingers in biochemistry busy for years. Several 1984 papers suggest that it has been a vintage year.

It is plainly necessary to understand the self-association of actin before one can hope to attain a proper grasp of the way in which it is affected by the many varieties of acting-binding protein that are present in the cell. An important early discovery was the fact that actin monomers bind ATP, which is hydrolysed in the course of polymerization of G-actin to long filaments.

The ADP resulting from hydrolysis remains bound to the actin molecules in the filament and inorganic phosphate is released into solution. The wider importance of this process became clear when it was discovered that microtubules assemble by an analogous mechanism that makes use of GTP in place of ATP. Many studies have focussed on the effect of ATP and ADP on actin polymerization and it has become clear that it is promoted by ATP. When only ADP is present on the actin monomers, polymerization is slower and the proportion of the free monomers left over at the end (in equilibrium with the filaments) is greater. Moreover, the process of ATP hydrolysis can allow actin filaments to 'treadmill', that is to lengthen at one end by incorporation of monomers while shortening at the other by release of subunits. This steady-state flux is caused by the difference of the affinity of the two filament ends for monomers. The ends are distinguished by the structural polarity of the filament and are called 'barbed' and 'pointed' ends because of their appearance in electron micrographs after labelling with myosin subfragments. The barbed ends have a higher affinity for actin monomers and are faster growing than the pointed ends.

Recently, there has been striking progress in understanding the fine details of actin polymerization. Rates of elongation of the two ends of filaments have been individually determined and the time course of ATP hydrolysis has been measured accurately and compared with the rate of elongation. Emerging from the combination of all these data, important facts about the dynamics of actin filaments have led to interesting conjectures about the function of actin in the cytoskeleton.

In particular it has been discovered that polymerization and ATP hydrolysis are not coupled^{1,2}. Actin monomers bearing ATP associate with the filament ends; hydrolysis follows about ten seconds later. Under conditions of polymerization, many new monomers can bind at the two ends during the period required for ATP hydrolysis. Thus, the ends of actin filaments carry a string of subunits with bound ATP, referred to as an 'ATP cap'³. Subunits near the middle of the filament possess bound ADP, since they are the first to assemble and have had sufficient time to allow ATP hydrolysis. Under conditions of depolymerization the composition of the filament ends is different. Filaments lose their ATP caps with the departure of the terminal subunits. Depolymerizing filaments therefore have ADP-containing subunits at their ends. Thus a consequence of the decoupling of polymerization and ATP hydrolysis is that the ends of the growing and shortening filaments differ in composition⁴.

If the identity of the nucleotide on the terminal subunits affects the rate of attachment and dissociation at these points, one can imagine readily that the presence or absence of ATP caps will have drastic consequences for assembly and disassembly of

filaments. The rates in question have been determined⁵; ADP-bearing subunits dissociate five times as fast from filament ends as ATP-subunits. A depolymerizing filament with terminal ADP-subunits loses monomers at five times the rate of a lengthening one with an ATP cap. In other words, the composition of a shortening filament end favours depolymerization and that of a lengthening filament end disfavors it. It follows that the preferred direction of filament growth depends on the concentration of actin monomers. At high monomer concentration, attachment of monomers prevails over dissociation, leading to lengthening filaments with ATP caps. At low monomer concentration, association is slower than dissociation.

Filaments thus shorten and possess terminal ADP-subunits. Near the so-called critical concentration, at which the transition from shortening filaments into lengthening filaments occurs, relatively small changes in the monomer concentration lead to a major change in the polymerization mechanism of actin. Just below the critical monomer concentration, ADP-subunits dissociate rapidly from filament ends; just above it, filaments develop ATP caps and dissociation of the terminal ATP-subunits is slow. The lesson is that small changes in monomer concentration, under the right circumstances, can exert a greatly amplified effect on the composition of filament ends, very probably on their reactivity towards 'capping' proteins, and certainly on the growth rate of filaments⁴.

Despite this considerable advance in our understanding of actin assembly, important questions remain to be answered. For example, do ATP-monomers bind to an ATP cap at the same rate as to filament ends with ADP-bearing subunits? A recent theoretical discussion analyses the consequences that the answer might have⁶. If the ATP-monomers prefer ATP caps, lengthening filaments with their ATP caps will continue to grow because they can incorporate monomers rapidly, whereas shortening filaments, with terminal ADP-subunits, will go on depolymerizing because they bind monomers only slowly. Thus both lengthening and shortening fila-

ments will coexist in the solution, and will interconvert only slowly. No experimental evidence for this phenomenon yet exists, but the theoretical treatment shows what ramifications of the ATP-hydrolysing system may come to light.

Another open question is whether or not the two ends of the filaments behave differently in the above respects. Under physiological conditions, the barbed end binds and releases actin molecules about ten times faster than the pointed end. In most types of experiment, the sum of the growth rates at the two ends is measured. Because of the much higher activity at the barbed end, the information that emerges reflects mainly the events at that end, for they swamp the slower reactions occurring at the pointed ends. Many proteins are known to bind to the barbed ends and there modulate the polymerization and disassembly of

actin⁷. Indeed, only one cytoskeletal protein has so far been found to associate with the pointed ends of filaments⁸. Perhaps, in general, the barbed filament ends in cells are occupied by proteins which regulate their dynamics, whereas the pointed ends are most often free. This makes it all the more important to develop methods for investigating what goes on at the pointed ends of the filaments. □

1. Pardee, J.D. & Spudis, J.A. *J. Cell Biol.* **93**, 648 (1982).
2. Pollard, T.D. & Weeds, A.G. *FEBS Lett.* **170**, 94 (1984).
3. Brenner, S.L. & Korn, E.D. *J. biol. Chem.* **258**, 5013 (1983).
4. Carlier, M.-F., Pantaloni, D. & Korn, E.D. *J. biol. Chem.* **259**, 9983 (1984).
5. Pollard, T.D. *J. Cell Biol.* **99**, 769 (1984).
6. Chen, Y. & Hill, T.L. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7520 (1984).
7. Pollard, T.D. *Nature News and Views* **312**, 403 (1984).
8. Southwick, F.S. & Hartwig, J.H. *Nature* **297**, 303 (1982).

Albrecht Wegner is in the Institut für Physiologische Chemie, Ruhr Universität Bochum, Postfach 10 21 48, 4630 Bochum 1, FRG.

Drosophila development

Abdominal gene organization

from Phil Ingham

OVER the past few months, the growing interest in the developmental genetics of the fruitfly, *Drosophila melanogaster*, has received an added stimulus from the discovery of a 'homoeo box' sequence that characterizes some of the genes underlying development^{1,2}. One focus has been on the bithorax complex — a cluster of genes that control the development of all segments of the *Drosophila* body posterior to the second thoracic segment. Yet, while the discovery of several homoeo box sequences in the bithorax complex has facilitated identification and analysis of a transcript deriving from its distal or 'abdominal' region¹, understanding of this region at the developmental and genetic level has remained less than comprehensive. An extensive analysis by Morata and co-workers, published on page 108 of this issue³, goes a long way towards redressing the balance. Intriguingly, the genetic organization of the bithorax complex that emerges corresponds well with the distribution of homoeo box sequences (see ref. 4) within the complex.

The induction and identification of mutant alleles of genes has had an important impact on the study of the bithorax complex. Mutations mapping to the proximal region of the complex have dramatic effects on the development of the adult fly; *bithorax* (*bx*) and *postbithorax* (*pbx*), for example, transform the small metathoracic appendage, the haltere, into the much larger mesothoracic wing^{5,6}. Equally as striking is the replacement of part of the first abdominal segment by thoracic legs that is caused by *bithoraxoid* (*bxd*) mutations. Using these gross changes in developmental potential as an assay, E.B. Lewis has carried out a detailed dissection of the proximal region of the complex^{6,7,8}.

The complexity of the genetic organization that is revealed has since been confirmed in almost every particular at the molecular level⁹. Lewis has demonstrated the existence of four genes or functional units, *abx* (*anterobithorax*), *bx*, *pbx* and *bxd*, each of which controls the development of a specific developmental compartment¹⁰ from the posterior compartment of the mesothoracic segment to the anterior compartment of the first abdominal segment. All four genes seem to be subsumed by a fifth gene, named *Ultrabithorax* (*Ubx*), mutations of which fail to complement *abx*, *bx*, *pbx* and *bxd* mutations. The region of the body affected by all these mutations is thus termed the *Ubx* domain³. Molecular analysis has revealed that *abx*, *Ubx* and *bx* mutations all fall within a single large transcription unit (termed the *Ubx* unit) whilst *bxd* and *pbx* mutations map in a smaller adjacent transcription unit (the *bxd* unit)⁹. This genetic and structural complexity is still not fully understood, but it seems that transcripts from the *Ubx* unit are required in all four compartments that constitute the *Ubx* domain¹¹, and that transcription may be potentiated in two of the compartments by the *bxd* unit^{4,12}.

In contrast to these complexities of control by the proximal region of the bithorax complex, Lewis suggested that the development of each abdominal segment of the fly is simply controlled by one of the *infra-abdominal* (*iab*) genes⁶. Whilst this provides a reasonable working model of the bithorax complex as a whole, it makes no predictions about the nature (if any) of the hierarchical organization of the abdominal region of the complex. The novel and important finding of Morata and collaborators³ is that the abdominal region

100 years ago

ACCORDING to the *North China Herald* there died a few months ago at Peking, the greatest Chinese mathematician of the present century. His name was Li Shan-lan, and he was Professor of Mathematics at the Foreign College in the Chinese Capital. He differed from the mathematicians of Europe in this respect, that he denied the non-existence of a point. "A point," said Prof. Li, "is an infinitesimally small cube," and in saying this he only reproduced the theories of Chinese sophists 2000 years ago. A great writer of that age said: Subtlety is the occult part of the minute. Be a thing subtle or gross, it seems to me that it must have a form. But I take it that what is neither gross nor subtle can neither be talked of nor imagined.

From *Nature*, **31**, 227; 8 January 1885.

is further subdivided into two major genes which seem to be structurally and functionally similar to the *Ubx* gene.

Study of the abdominal region is hampered by the limited diversification between abdominal segments; segments A2 to A6 of the adult female are so similar that transformation of one to another would probably go undetected in a standard screen for mutants. Whereas dominant 'gain-of-function' mutations of abdominal genes have been recognized by their ability, for example, to transform the last thoracic segment to an abdominal segment or cause a male-specific A5 pigmentation pattern in A4, and some 'loss-of-function' mutations have been induced by reverting these dominant mutations, the *iab* genes have in large part remained hypothetical. Morata and co-workers have circumvented the problem of identifying abdominal mutants by employing a screening protocol that allows recognition of mutations in the ultrathorax complex solely on the basis of their lethality (non-lethal mutations causing an easily-visible phenotype can also be picked). Whilst this may seem an obvious expedient, it is worth remembering that none of the *Ubx* subfunctions mutate to lethality and so the method is not guaranteed to identify all elements of the complex. Nor indeed have Morata and collaborators succeeded in isolating mutations corresponding to each postulated *iab* locus. But they have been able to identify two major genes, designated *abdominal-A* (*abd-A*) and *Abdominal-B* (*Abd-B*), which together control the development of all segments posterior to the *Ubx* domain. In both cases, and like the *Ubx* domain, the realms of action of these genes extend over more than one segment.

Moreover, the *abd-A* domain begins not at a segment border but rather at the boundary between anterior and posterior compartments in segment A1 (where the *Ubx* domain ends¹³) thus confirming that bithorax complex genes act on compartmental rather than segmental units^{13,14}. Mutations in *Abd-B* affect the development of segments A5 to A8 but have no effect on more anterior segments; the *Abd-B* gene therefore subsumes the functions previously attributed to the *iab-5, 6, 7* and *8* genes. Similarly, *abd-A* apparently subsumes the functions of *iab-2, 3* and *4* since *abd-A* mutations transform segments A2, A3 and A4 to A1. This is demonstrable in the case of *iab-2* because a mutant allele of this gene already exists¹⁵. As expected, in *iab-2/abd-A* animals, segment A2 is transformed to A1 whilst A3 and A4 are unaffected. Thus *iab-2* appears to represent a sub-function of the *abd-A* gene, in a way comparable to the relation between *bx* and *Ubx*. The analogy will be considerably strengthened if mutations in the elusive *iab-3* and *iab-4* genes also turn out to be allelic to *abd-A*.

The results of this study have considerable significance for the future analysis of the bithorax complex. They provide the

comprehensive description of the structural and functional organization of the abdominal genes that will be essential to a full understanding of the region at the molecular level. One obvious inference is that the abdominal region contains two further major transcription units, each analogous to the *Ubx* unit. This prediction had been foreshadowed by the mapping of two homoeo boxes within the complex but distal to *Ubx*^{1,4}. Whether the transcripts deriving from the *abd* genes are modulated in different compartments by *iab* analogues of the *bx* unit remains an open question. The coming months should prove to be as exciting and revealing as the past year has been. □

Dating techniques

Lighting up the past with lasers

from Ann G. Wintle

THE same basic property of common minerals, such as quartz and feldspar, that enables pottery and sediments to be dated by thermoluminescence has now been exploited in a new optical dating method, described by D.J. Huntley, D.I. Godfrey-Smith and M.L.W. Thewalt on page 105 of this issue. The build up of electrons trapped at defects in crystals of the minerals when they are exposed to ionizing radiation is the basis of both techniques. Provided that the electrons already trapped in the minerals had been removed in the firing of pottery or the deposition of sediment, then the electrons accumulated thereafter and experimentally untrapped by thermal or optical means will provide a measure of the sample's age.

The new optical-dating method has been applied to quartz grains from several sedimentary environments. The trapped electrons which have built up in the grains since they were last exposed to light at deposition are optically excited from their traps by the 514.5 nm beam from an argon-ion laser. A portion of these electrons recombine within the crystals and produce luminescence. The emission spectra of naturally-occurring quartzes are complex and vary with temperature, the origin of the samples and their previous exposure to radiation (McKeever, S.W.S. *Rad. Protect. Dosimetry* 8, 81; 1984). At room temperature the spectra of most samples are dominated by a broad blue emission in the region 450–470 nm. Using suitable filters both for the argon beam and in the detection system, the luminescence in this region can be detected with a high signal-to-noise ratio whilst the sample is being optically excited.

This development has come at a particularly interesting time in the application of thermoluminescence dating to sediments. In Europe, most dating has been carried out on the feldspar grains from loess thought to be less than 130,000 years old

1. McGinnis, W. *et al.* *Nature* 308, 428 (1984).
2. Scott, M.P. & Weiner, A.J. *Proc. natn. Acad. Sci. U.S.A.* 81, 415 (1984).
3. Sánchez-Herrero, E., Vernos, I., Marco, R. & Morata, G. *Nature* 313, 108 (1985).
4. North, G. *Nature News and Views* 311, 214 (1984).
5. Lewis, E.B. *Am. Zool.* 3, 33 (1963).
6. Lewis, E.B. *Nature* 276, 565 (1978).
7. Lewis, E.B. in *Developmental Biology Using Purified Genes* (eds Brown, D.D. & Fox, C.F.) 189 (Academic, New York, 1981).
8. Lewis, E.B. in *Embryonic Development: Genes and Cells* (ed. Burger, M.) 269 (Liss, New York, 1982).
9. Bender, W. *et al.* *Science* 221, 23 (1983).
10. García-Bellido, A., Ripoll, P. & Morata, G. *Nature* 245, 251 (1973).
11. Akam, M.E. *EMBO J.* 2, 2075 (1983).
12. Ingham, P.W. *Cell* 37, 815 (1984).
13. Struhl, G. *Nature* 308, 454 (1984).
14. Hayes, P.H., Sato, T. & Denell, R.E. *Proc. natn. Acad. Sci. U.S.A.* 81, 545 (1984).
15. Kuhn, D.T., Woods, D.F. & Cook, J.L. *Mol. gen. Genet.* 181, 82 (1981).

Phil Ingham is at the Imperial Cancer Research Fund Laboratories, London NW7 1AD, UK.

(Wintle, A.G., Shackleton, N.J. & Lautridon, J.P. *Nature* 310, 491; 1984). But it has recently been shown that the thermoluminescence signal from feldspar is unstable, cannot be used at all for samples older than about 120,000 years (Debenham, N.C. *Nuclear Tracks*; in press) and would result in underestimating by greater than 10 per cent the age of samples that are older than 30,000 years.

This problem will result in a move to quartz as the more suitable mineral for thermoluminescence dating, despite the fact that it has a lower thermoluminescence sensitivity and is less easily bleached than feldspar. Because quartz loses its trapped electrons relatively slowly on exposure to light, there is always the worry that extended laboratory bleaching, either with a sunlamp or sunlight, may not give the correct assessment of the trapped electron population at the time of deposition. This is most likely to be true for glacial melt-water sands or river sediments. If the laboratory bleach is too long, then the signal obtained by thermoluminescence measurements will be too large and result in an overestimate of age.

The new optical dating method avoids this problem, because the signal observed immediately after exposure to the laser beam is that from the most light-sensitive electron traps, which would also have been the first to be emptied at the time of deposition. These electron traps appear to be much more light sensitive than those on which thermoluminescence measurements depend; a ten-second exposure to sunlight severely depletes them of electrons. The new method should therefore be a means of dating water-lain deposits and glacial sediments, which are likely to have had only a short exposure to light. □

Ann G. Wintle is at the Godwin Laboratory, Sub-department of Quaternary Research, University of Cambridge, CB2 3RS, UK.

Semiconductor technology

Kinetics of pulsed laser annealing

from Ian W. Boyd and Steven C. Moss

PULSED laser annealing (PLA) is a technique which uses intense bursts of laser radiation to repair the damage produced by ion implantation of crystalline semiconductors, thus improving the structural order and the electrical properties of the annealed material. The mechanism behind the technique has been the subject of a great deal of heated debate over the past few years^{1,2}, but evidence compiled from a wide variety of measurements now seems to favour strongly a thermal model over a non-thermal model.

Initially, microscopic examination of the irradiated surface of the semiconductor and evidence for the redistribution of impurity atoms within the material suggested that it had melted and then recrystallized by liquid-phase epitaxial regrowth. In that case, the laser radiation was thought to excite electrons in the valence band across the bandgap and into the conduction band, from which they very rapidly relaxed back to the valence band, transferring their energy to the lattice, and thereby heating up and melting the material. Alternatively, it was suggested that so many electrons could be excited into the conduction band that the usually strong covalent bonding of the solid would be severely weakened. In that case, it was possible to envisage a phase transition whereby, instead of violent atomic vibrations breaking the solid into a hot liquid phase in the usual thermal melting sense, the material 'wobbled' itself gently into a cool jelly-like structure.

There followed a period of unprecedented activity in numerous laboratories to differentiate between the two models. Unfortunately, many experiments were performed too hastily, some were extremely difficult, and others were extremely difficult to interpret. Early results favoured the non-thermal model, which was vague enough to accommodate many strange phenomena. However, over the following years, experiments were repeated, reanalysed and sometimes reinterpreted. On the basis of experiments centred around the measurement of the lattice temperature during and after the annealing pulse, an onslaught on the non-thermal model was mounted. The structure and physical properties of the irradiated material were also subjected to unique and most ingenious investigations^{1,2}.

Early support for the thermal model was strongly initiated by time-resolved reflectivity measurements on a nanosecond scale, using a weak probe beam centred on the irradiated layer during the annealing event³. The observed high reflectivity phase was characteristic of liquid silicon and transient transmission measurements in the same regime have recently been

shown to be consistent with melting⁴. Previous measurements of the same kind, which had appeared to favour non-thermal mechanisms, were apparently in error, because of detector artefacts. Time-of-flight mass spectrometry, in conjunction with transient reflectivity⁵, enabled an analysis of the atoms ejected from the surface during the annealing process. Early measurements provided some support for the non-thermal model by showing that melting occurred only for pulses more than 75 per cent greater in energy than the previously determined annealing threshold, but a careful repeat of the experiment has since shown that the evaporation rate at the PLA threshold is indicative of temperatures exceeding the melting point⁵. Independent photoemission studies (see R. T. Williams *et al.* in ref. 2), as well as time-resolved structural measurements with synchrotron X rays⁶ and low energy electron diffraction⁷, are also in agreement with the thermal-melting model. Transient electrical conductance of the irradiated material not only yields strong evidence of rapid melting but also provides quantitative estimates for melting and solidification velocities⁸.

Great controversy has surrounded the interpretation of several sets of Raman scattering data that have been recorded at particular frequencies during the laser annealing cycle. The earlier work was seen as strong evidence against thermal melting, since maximum lattice temperatures around 600 K were deduced (silicon melts at 1685 K). More recently, frequency-resolved Raman data taken at fluences just below the transition threshold (the high reflectivity phase cannot be Raman active) have indicated that temperatures close to the melting temperature of silicon can be reached. Furthermore, temporal resolution artefacts during the kind of high energy PLA used in earlier measurements result in misleadingly low estimates of the maximum temperature (see Von der Linde *et al.* in ref. 1).

Clearly, in the nanosecond excitation regime, high lattice temperatures and melting occur in extremely short time periods. But how long does it take for the energy in the highly excited electronic plasma to be transferred to the lattice? Picosecond studies of laser-induced phase transitions have been most useful in elucidating the various optical- and electronic-scattering phenomena. Using 15–20 ps pulses, and monitoring the reflectivity and transmission of a weak probe beam, it has been deduced that lattice melting occurs during the pulse duration (ref. 9, and Bucksbaum & Baker in ref. 1) and that the maximum electron-hole plasma concent-

ration during the 20 ps pulses was less than 10^{21} cm^{-3} (ref. 9). Also, by observing the thermally-induced variation of the multiple reflections in silicon-on-sapphire during annealing, direct measurement of lattice heating has been possible (see Lompre *et al.* in ref. 1). An analysis of the emission of charged particles from the sample surface is also consistent with thermal melting, sets an upper limit of 5000 K on the electronic temperature and suggests that thermalization between carriers and lattice occurs within 10 ps (see Malvezzi *et al.* in ref 1).

In the course of rapid melting, it is possible to obtain such fast cooling rates that the liquid layer freezes in an amorphous configuration before recrystallization can occur¹⁰. Such 'amorphization' is clearly consistent with thermal arguments. Optically-induced phase transitions have also been studied during the past year using 55–90 fs pulses. Time-resolved reflectivity clearly shows that the energy transfer from the excited carriers to the crystal lattice, and subsequent melting, occurs within a few picoseconds¹¹. An independent study of the structural changes, in which the symmetry dependence of the second-harmonic generation from a weak probe pulse reflected from the crystal was monitored, clearly indicated a transition from order to disorder on a time scale less than a picosecond, again consistent with rapid melting¹². However, an analysis of recent measurements of the self-induced reflectivity of intense 100 fs pulses incident upon silicon requires carrier densities very close to the total number of valence electrons, and energy relaxation rates that are 300 times larger than the low density relaxation rates¹³. If confirmed, these results may once again imply the necessity of a non-thermal role in the kinetics of laser-induced 'melting' on the femtosecond time scale. On this time scale, however, the photo-excited system is far from equilibrium, and the definitions applicable to longer time scales become less appropriate to describe any of the transient states of the matter. Until the appropriate measurements are performed and until physical properties are redefined in this regime, the now well worn and metamorphic non-thermal model must remain on the desks of the inventors. □

1. *Energy Beam-Solid Interactions & Transient Thermal Processing* (eds Fan, J.C.C. & Johnson, N.M.), North-Holland, 1984).

2. *Laser and Electron-Beam Interactions with Solids* (eds Appleton, B.R., & Celler, G.K.), North-Holland, 1982).

3. Austin, D.H. *et al.* *Appl. Phys. Lett.* **33**, 437 (1978).

4. Lowndes, D.H. *Phys. Rev. Lett.* **48**, 267 (1982).

5. Pospiesznyk, A. *et al.* *J. appl. Phys.* **54**, 3176 (1983).

6. Larsen, B.C. *et al.* *Appl. Phys. Lett.* **42**, 282 (1983).

7. Becker, R.S. *et al.* *Phys. Rev. Lett.* **52**, 307 (1984).

8. Galvin, G.J. *et al.* *Phys. Rev. Lett.* **48**, 33 (1982).

9. Liu, J.M. *et al.* *Appl. Phys. Lett.* **41**, 643 (1982).

10. Boyd, I.W. *et al.* *Appl. Phys. Lett.* **45**, 80 (1984).

11. Shank, C.V. *et al.* *Phys. Rev. Lett.* **50**, 454 (1983).

12. Shank, C.V. *et al.* *Phys. Rev. Lett.* **51**, 900 (1983).

13. Hulin, D. *et al.* *Phys. Rev. Lett.* **52**, 1998 (1984).

Ian W. Boyd is in the Department of Electronic and Electrical Engineering, University College, London WC1E 6BT, UK; Steven C. Moss is at Aerospace Corporation, PO Box 92957, Los Angeles, California 90009, USA.

Construction of linkage maps with DNA markers for human chromosomes

Ray White, Mark Leppert, D. Timothy Bishop*, David Barker†, Jon Berkowitz, Candace Brown, Patricia Callahan, Tom Holm & Leslie Jerominski

Howard Hughes Medical Institute and Departments of Cellular, Viral and Molecular Biology, and

* Department of Medical Biophysics and Computing, University of Utah Medical School, Salt Lake City, Utah 84132, USA

DNA markers and sampling of three-generation families can be used to construct complete linkage maps of human chromosomes. This is important in mapping disease loci and in determining the genetic or environmental component of a disease.

A RELATIVELY new system of genetic markers, based on direct detection of DNA sequence polymorphism with restriction enzymes, has made it possible to initiate the construction of a detailed genetic linkage map for the human¹. The new genetic loci are defined by cloned DNA segments which can represent genes of known specificity or arbitrary loci without known function. In addition to their localization to specific regions of specific chromosomes by physical methods^{2,3}, the linkage distances and gene orders which characterize these new markers can be estimated by family studies.

The human genome is estimated to span 3,300 centimorgans; localization of a new locus can be achieved in studies of reasonable scale if marker loci are spaced at 40 centimorgans⁴. It follows that only 80–100 evenly-spaced genetic marker loci may be needed to cover the human genome and allow the placement of any new marker with a high degree of reliability. However, in order to obtain evenly spaced loci, a much larger number of marker loci must first be placed on the map^{5,6}. This will be a large-scale endeavour and a high efficiency of data collection will be important.

Applications of markers

A linkage map with its constituent markers will contribute to the localization of a number of human genetic diseases, for example cystic fibrosis, on the human gene map. As most human genetic diseases have no known phenotype in cultured cells, they cannot be localized to specific chromosomes using hybrid cell mapping approaches. Their position on the gene map can, however, be determined by cosegregation in families with linked genetic markers.

A set of such markers derived from a complete map has the advantages of even spacing, therefore markers from the same regions are not tested redundantly, with at least two linking to any genetic disease locus. Furthermore, because the genetic distances between the markers are known accurately, increased analytical resolution in mapping genetic disease loci will be possible. Such precision will be essential in smaller-sized families where genetic data are limited.

When genetic disease loci are mapped, the adjacent polymorphic DNA markers serve as preclinical indicators of genotypic status for individuals at risk. Duchenne⁷ and Becker⁸ muscular dystrophy as well as Huntington's disease⁹, X-linked retinitis pigmentosa¹⁰ and fragile X-linked mental retardation¹¹ have been linked already with DNA markers.

A second important application is the genetic characterization of common traits and diseases and the determination of the genetic or environmental basis of disorders which cluster in families, for example, breast cancer^{12,13}, schizophrenia¹⁴ and some forms of epilepsy¹⁵. If individuals of a specific genotype

only are susceptible to the disease, the disorder has a significant genetic risk factor. Cosegregation of the disorder with a genetic marker within the family can provide this demonstration. Not only those disorders which have a genetic component can then be discovered, but also individuals at risk within such families can be identified.

Over 200 polymorphic DNA markers have been reported¹⁶. Although most of the reported markers have only two alleles, a few have multiple alleles, apparently because of single base changes affecting restriction sites. As frequent heterozygosity is a critical requirement for linkage studies, the number and frequency of alleles available at a polymorphic locus characterize its usefulness as a genetic marker. Multi-allelic human loci can result from multiple, independent restriction site alterations

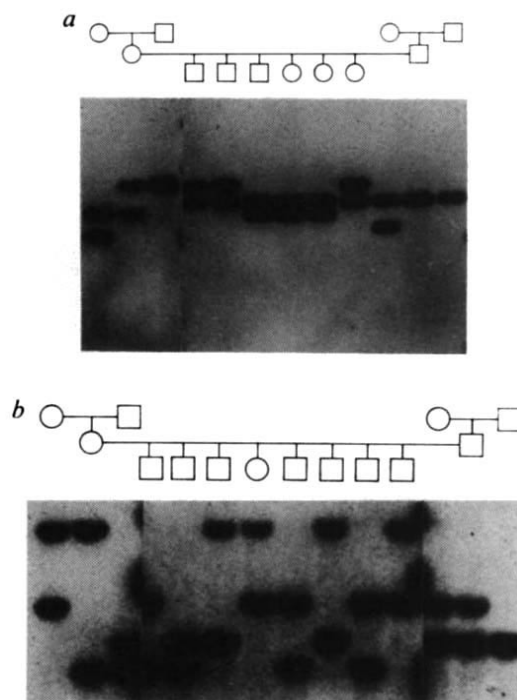


Fig. 1 *a*, Autoradiogram of a Southern blot transfer from a three-generation family. Each lane contains DNA from the individual represented in the pedigree directly above. The DNA was cut with the restriction enzyme *PvuII* and hybridized with the insulin probe phins 310. Alleles in this particular family range in size from 0.9–1.1 kilobases (kb). *b*, Autoradiogram of a Southern transfer from a three-generation family probed with the Harvey *ras-1* gene. The human DNA was cut with the restriction enzyme *TaqI* and the alleles observed range in size from 4.2 to 2.0 kb. The genotypes of the individuals are shown directly below their place in the pedigree.

† Present address: Collaborative Research, Inc., 128 Spring Street, Lexington, Massachusetts 02173, USA.

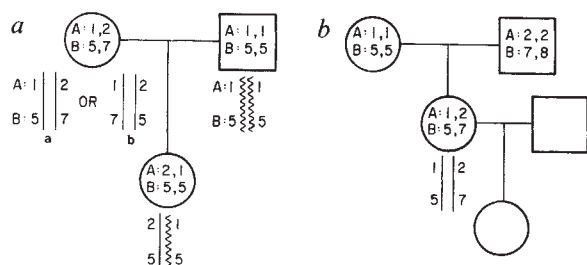


Fig. 2 Ambiguous determination of phase at two loci, A and B, when only parental genotypes are known. *a*, The two possibilities for phase in the mother are shown as *a* and *b*. *b*, Genotypes of both grandparents are known, thus phase in a parent may be determined with certainty.

at the locus, such as those characterized for the β -globin genes¹⁷. Useful multi-allelic loci have been defined also in regions which include sets of short tandem repeats varying in number in the population, such as those characterized at the insulin locus¹⁸.

A set of genetic markers localized to the distal one-third of the short arm of chromosome 11 by independent methods has been used for an initial study of linkage among DNA markers¹⁹. These marker loci have multiple alleles; for example, the insulin gene locus¹⁸ and the Harvey *ras-1* gene locus^{20,21} each reveal several alleles (Fig. 1). These DNA polymorphisms are due to a series of deletions in a set of short tandem repeats at the locus^{18,22}. Two other loci located in the same region of chromosome 11p, the β -haemoglobin gene region¹⁷ and an arbitrary locus called D11S12 (ref. 23), each reveal several polymorphic restriction sites. As the restriction sites are so closely spaced, the resulting haplotypes will be disrupted by recombination only very infrequently and can be treated as alleles. The multi-allelic character of this set of marker loci makes them very useful for recombination studies since a high proportion of meioses occur with multiply heterozygous chromosomes which allow the detection of recombination.

Linkage

Determination of a heterozygous parent genotype does not reveal the distribution of alleles between the two chromosomes. These distributions influence the interpretation of results because a given progeny chromosome will be scored as either recombinant or non-recombinant depending on the arrangement chosen (Fig. 2*a*). Because of the analytical complexities resulting from such incomplete knowledge of the genotypes of individuals in families, linkage data from humans have traditionally been evaluated by likelihood analysis²⁴, where an estimate of the recombination interval (r) between two loci is obtained by determining the value of r which gives the maximum probability for the observed data. As one of the traditional problems in human genetics is the determination of whether two loci are linked, a quantitative expression for the likelihood of linkage is the ratio of the probability for the data set at the maximum likelihood value of r to the probability at $r=0.5$ (the marker loci are genetically unlinked). This ratio is usually expressed as a logarithm, the LOD (log of the odds) score for linkage. LOD scores of three or higher, representing a relative likelihood of linkage to non-linkage $\geq 1,000:1$ are considered to be strong evidence for linkage. (On average, the LOD score at any specified value of r is proportional to the number of informative chromosomes in the study; that is, the number of chromosomes which can be scored as either recombinant or non-recombinant.)

Extended, multigenerational kindreds have been considered as the most effective structure for human linkage studies^{25,26}. The multigenerational character of such families can sometimes provide phase information in parents because determination of grandparental genotypes will unambiguously reveal the allelic distribution (allele phase) in the parents unless both grand-

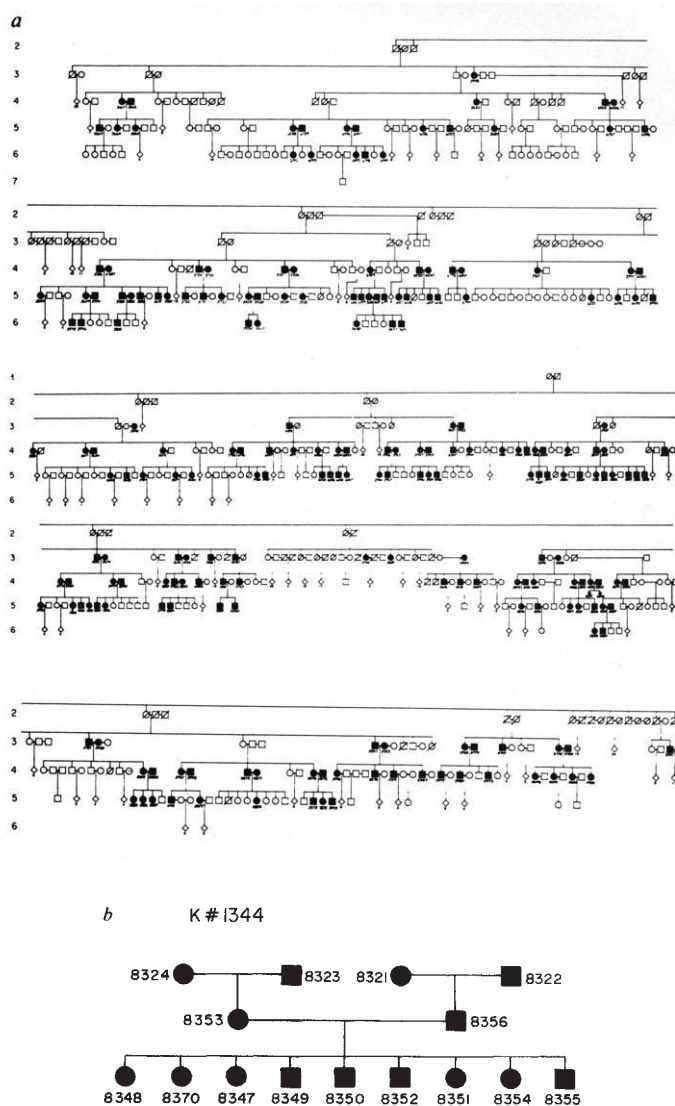


Fig. 3 *a*, Kindred 1085, an extended multigeneration family initially used for linkage studies on 11p. Individuals actually sampled are designated by the filled symbols. *b*, Pedigree of sampled individuals (filled squares and circles) in a typical three-generation, large sibship-size family.

parents are identical heterozygotes (Fig. 2*b*). Importantly for studies of inherited diseases, use of a single family also minimizes the possibility that genes at more than one locus might be involved.

Large sibship sizes are needed for the most efficient data collection. They yield more informative chromosomes associated with the constant amount of work required for the determination of the parents' and grandparents' genotypes and also provide a strong basis for a probabilistic estimate of the parental allelic distribution (phase). This becomes important when the allele distribution is not explicitly defined by the genotypes of the grandparents. We determined genotypes at the four loci for some 206 individuals within a single kindred, K1085 (ref. 27).

To obtain a set of samples representing larger sibships as well as grandparents, a number of large and complete three generation families were ascertained anecdotally. We sampled 18 families each with four living grandparents and seven or more children over five years old, as well as ten families with similar characteristics. One example is shown in Fig. 3. Genotypic data at the four marker loci, A, B, H and I, were obtained for 367 individuals from the 28 families, and gene order on the locus was analysed.

Table 1 Maximum likelihood estimates of recombination fractions and LOD scores for linkage between four loci on the short arm of chromosome 11

Combined analysis				
	B	A	I	H
B	-	72	49	25
A	2.9±0.9	-	57	27
I	11.6±1.6	9.1±1.5	-	59
H	12.5±2.3	10.3±2.2	3.2±1.1	-
Analysis of K1085				
	B	A	I	H
B	-	11	6.7	3.8
A	3.2±2.3	-	8.8	4.6
I	15.1±3.8	11.6±3.2	-	9.9
H	14.4±5.2	10.5±4.6	6.7±2.9	-
Complete, three-generation families				
	B	A	I	H
B	-	61	43	21
A	2.9±1.0	-	48	22
I	10.6±1.8	8.1±1.6	-	50
H	12.0±2.5	10.2±2.5	1.9±1.0	-

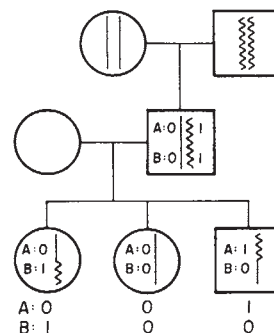
The combined linkage analysis is for both the extended kindred 1085 and the set of 28 large, three-generation families. Percentage recombination±s.e.m. is shown below the diagonal and the associated LOD scores above the diagonal. Linkage analysis for K1085 alone and for the complete, three-generation families alone are shown also. B, β -Haemoglobin gene locus; A, D11S12; I, insulin gene locus; H, Harvey *ras-1* oncogene locus. In all cases, analysis was performed with PAP³².

The maximum likelihood estimates of the recombination fractions and associated LOD scores for each of the six intervals defined by the four marker loci are shown in the combined analysis of Table 1. The order B-A-I is indicated by the estimate of 12% recombination between B and I which is larger than either the observed 3% between B and A or the 9% between A and I. Similarly, the order A-I-H is suggested by the estimated A-H distance of 10%, larger than the 9% A-I interval or the 3% I-H interval. Combining these two gene orders for the loci taken three at a time suggests the order B-A-I-H for the four loci; this result is supported by the observed 13% recombination between B and H. The pairwise LOD scores for linkage derived from this data set were 25-72. Although approximately equal numbers of male and female meioses were observed, the majority of the recombinants derive from male meioses.

Linkage is firmly established with the high LOD scores reported here. The maximum likelihood recombination fraction estimates and associated LOD scores for the extended kindred (K1085) and the complete three-generation families are shown in Table 1. Although we observed some differences in the estimates of recombination fractions in the two sets of families, only one difference is greater than one standard deviation. These values for the intervals defined by two factor crosses also agree reasonably well with previous studies^{28,29}.

An estimate of efficiency of the several family structures was obtained by calculation of the average LOD score per individual sampled, which yields an average of 0.04 LOD per individual over all four loci in ref. 29 and the K1085 data reported here (Table 1). In contrast, an average of 0.11 LOD per individual is obtained with the complete, three-generation families (Table 1), which we ascribe to the specifically ascertained structural features of sibship size and sampled grandparents. Clearly, these highly efficient complete three-generation families are an effective resource for the continued development of the human linkage map.

With our data set it is convenient also to examine the relative recombination frequency between males and females. It has been reported that the recombination frequency is greater in females than in males³⁰. In this study, however, we see the

**Fig. 4** Schematic diagram of the construction of the chromosome distribution pattern (CDP) for two loci, A and B. 0, Grandmaternal origin for a given locus; 1, grandpaternal origin.

contrary; of 41 recombinant chromosomes from parents in the complete three-generation families, 29 were derived from the male parent and 12 from the female parent.

We next examined the recombination history of specific chromosomes in the complete three-generation pedigrees. For the majority of parents heterozygous at any of the marker loci, we could determine the grandparent origin of the alleles. The numbers of recombinant and non-recombinant chromosomes could thus be counted for each pair of loci. The number of informative chromosomes was 47-228 over the intervals tested, yielding an actual efficiency (in terms of informative chromosomes per individual sampled) of 0.5-0.8. As we expected, the recombination frequencies were very close to those obtained by a complete likelihood analysis.

As genotypic data accumulate within these reference families, they will become more valuable for the primary chromosomal localization of new markers as well as for high-resolution mapping. As a large number of two-factor analyses would be slow, initial localization of a new locus can be achieved by comparison of the grandparent origin of the new locus with the grandparental origins of other loci previously characterized in the progeny chromosomes. The method is analogous to that used in mapping with recombinant inbred lines of mice. The notion is to assign a grandparent origin (either grandmother or grandfather) to the allele present at each locus of a progeny chromosome (Fig. 4). Each locus is characterized by a string of 0s (grandmaternal) and 1s (grandpaternal) with as many elements as there are informative progeny chromosomes, and referred to as the chromosome distribution pattern (CDP). Tightly-linked loci will have nearly identical chromosome distribution patterns and the frequency of differences will represent accurately the recombination frequency. A location for each new locus will be determined readily by comparison of the CDP of the new locus with the previously-determined CDPs in the data base (Table 2). The CDPs for each of the four loci on chromosome 11 are very similar, consistent with their tight linkage (Table 2). Comparisons of binary signatures of new loci with a large number of previously-characterized loci can be obtained very quickly.

Although we expect that estimates of recombination fractions will evolve continually with the acquisition of new data, it would be useful to have a map for which the gene order was not under constant revision (see ref. 31).

The most effective approach to the determination of gene order examines recombination among the offspring of individuals heterozygous for three or more loci, such that recombination in several intervals can be followed in the same set of chromosomes. This approach places a high premium on genetic marker systems of high quality, such that most individuals are heterozygous at several loci.

As there were a number of multiply informative meioses in the data set described previously, we performed an analysis of relative likelihood of gene order (Table 3). The logarithm of the relative likelihood of the gene order B-A-I-H over B-A-H-I

Table 2 Chromosome distribution patterns (CDP) for four 11p loci

Maternal chromosomes										Paternal chromosomes														
Chromo- some	B	A	I	H	Chromo- some	B	A	I	H	Chromo- some	B	A	I	H	Chromo- some	B	A	I	H	Chromo- some	B	A	I	H
2	1	1	1	-	110	-	1	1	1	179	0	0	0	-	278	-	1	1	1	384	1	1	-	-
3	0	0	0	-	111	-	1	1	1	180	1	1	1	-	279	-	0	0	0	385	1	1	-	-
4	1	1	1	-	112	-	1	1	1	181	1	1	1	-	280	-	1	1	1	386	1	1	-	-
5	0	0	0	-	113	-	0	0	0	182	0	0	0	-	281	-	0	0	0	387	0	0	-	-
6	0	0	0	-	114	0	0	0	-	183	1	1	1	-	282	-	0	0	0	388	0	0	-	-
7	0	0	0	-	115	0	0	0	-	184	1	1	1	-	283	-	1	1	1	389	1	1	-	-
19	0	0	0	-	116	0	0	0	-	188	-	0	0	-	295	0	0	0	0	390	0	0	0	0
20	0	1	1	-	117	0	0	0	-	190	-	1	1	-	296	1	1	1	1	391	0	0	0	0
21	0	0	0	-	118	0	0	0	-	191	-	1	1	-	297	1	1	0	0	392	1	1	1	1
39	1	1	1	-	119	0	0	0	-	192	-	0	0	-	315	1	1	1	1	393	1	1	1	1
40	1	1	1	-	121	1	-	1	-	193	-	1	1	-	316	1	1	1	1	394	0	0	0	0
41	0	0	0	-	122	0	-	0	-	194	-	1	1	-	317	1	1	1	1	395	1	1	1	1
42	1	1	1	-	123	0	-	0	-	195	-	0	0	-	318	1	1	0	0	397	-	1	1	1
43	0	0	0	-	125	1	-	1	-	196	-	0	0	-	319	0	0	0	0	398	-	0	0	0
44	1	1	1	-	126	1	-	1	-	197	-	1	1	-	320	0	0	0	0	399	-	0	0	0
45	1	1	1	-	127	1	-	1	-	198	-	0	0	-	321	1	1	1	1	401	-	1	0	0
47	1	-	0	-	128	1	-	1	-	214	-	-	1	-	323	0	-	0	-	402	-	1	1	1
48	0	-	-	-	129	-	0	0	0	215	-	-	1	-	324	1	-	-	-	403	-	1	1	1
49	1	-	-	-	130	-	0	0	0	216	-	-	1	-	325	0	-	-	-	404	-	0	1	1
50	1	-	1	1	131	-	1	1	1	217	-	-	1	-	326	0	-	0	0	405	-	1	-	1
51	1	-	1	1	132	-	0	0	0	218	-	-	0	-	327	0	-	0	0	406	-	1	-	1
52	0	-	0	0	133	-	0	0	0	220	-	-	0	-	328	1	-	1	1	407	-	0	-	0
57	0	0	0	-	134	-	0	0	0	221	-	-	0	-	333	-	-	0	0	408	-	0	-	0
58	1	1	1	-	135	-	0	0	0	222	-	-	1	-	334	-	-	1	1	409	-	0	-	0
59	0	0	0	-	136	-	0	0	0	223	-	-	0	-	335	-	-	0	0	410	-	0	-	0
60	0	0	0	-	137	-	0	0	0	225	-	-	1	-	336	-	-	0	0	411	-	0	-	0
61	0	0	0	-	138	-	0	-	0	226	-	-	0	-	337	-	-	0	0	412	-	0	-	1
62	0	0	1	-	139	-	1	-	1	227	-	-	0	-	338	-	-	0	0	413	-	1	-	1
63	0	0	1	-	140	-	0	-	0	228	-	-	1	-	339	-	-	0	0	414	-	1	1	1
67	0	0	0	-	141	-	0	-	0	229	-	-	0	-	343	-	0	0	0	415	-	1	1	1
68	0	0	0	-	142	-	1	-	1	230	-	-	1	-	344	-	1	1	1	416	-	0	0	0
69	0	0	0	-	143	-	0	-	0	231	-	-	0	-	345	-	1	1	1	417	-	0	0	0
70	0	0	0	-	144	-	0	-	0	232	-	-	1	-	346	-	0	0	0	418	-	1	1	1
71	0	0	0	-	145	-	1	1	1	233	-	-	1	-	347	-	0	0	0	419	-	0	0	0
72	0	0	0	-	146	-	1	1	1	234	-	-	0	-	348	-	1	1	1	420	-	0	0	0
74	-	-	0	0	147	-	1	1	1	235	-	-	1	-	350	0	0	0	-	429	0	0	0	-
75	-	-	0	0	148	-	0	0	0	236	-	-	0	-	351	1	1	1	-	430	0	0	0	-
76	-	-	1	1	150	-	0	0	0	241	-	-	0	-	352	0	0	0	-	431	1	1	1	-
77	-	-	0	0	151	-	0	0	0	242	-	0	1	-	353	1	1	1	-	432	1	0	0	-
78	-	-	0	0	152	-	1	1	1	243	-	-	0	-	354	1	1	1	-	433	0	0	0	-
79	-	-	0	0	153	0	0	0	-	245	-	1	1	-	355	0	0	0	-	434	0	1	1	-
80	-	-	0	0	154	1	1	1	-	246	-	0	0	-	356	0	0	0	-	435	0	0	0	-
81	-	-	1	1	155	1	1	1	-	247	-	1	1	-	357	0	0	0	-	436	1	1	1	-
82	-	-	0	0	156	1	1	1	-	248	-	-	1	-	358	1	1	0	-	437	0	0	0	-
86	-	1	1	-	157	1	1	1	-	249	-	-	0	-	362	-	1	1	1	439	-	1	1	-
88	-	-	1	-	158	1	1	1	-	266	-	-	0	0	364	-	-	0	0	441	-	0	0	-
89	-	-	1	-	159	1	1	1	-	267	-	-	1	1	365	-	-	0	0	442	-	0	0	-
90	-	-	1	-	160	1	1	1	-	268	-	-	1	1	366	-	-	0	0	443	-	0	0	-
91	-	-	0	-	161	0	0	0	-	269	-	-	0	0	367	-	-	0	0	444	-	1	1	-
92	-	-	1	-	163	-	-	1	1	270	-	-	0	0	368	-	-	0	0	445	-	0	0	-
93	-	0	0	-	165	-	-	1	1	271	-	-	1	1	369	-	0	0	0	447	0	0	0	-
94	-	1	1	-	166	-	-	1	1	273	-	-	0	0	370	-	1	1	0	448	0	0	0	-
95	-	0	0	-	167	-	-	0	0	274	-	-	1	1	371	-	0	0	1	449	1	1	1	-
98	-	1	1	1	168	-	-	0	0	275	-	-	0	0	374	-	-	0	0	450	0	0	0	-
99	-	1	1	1	169	-	-	0	0	276	-	-	1	1	375	-	-	1	1	451	0	0	1	-
100	-	0	0	0	171	-	-	1	-						376	-	-	1	1	452	0	0	0	-
101	-	0	0	0	172	-	-	0	-						377	-	-	0	0	454	1	-	1	-
102	-	1	1	1	173	-	-	1	-						378	-	-	0	0	455	0	-	0	-
105	-	0	0	0	174	-	-	1	-						379	-	-	1	1	456	0	-	0	-
107	-	1	1	1	175	-	-	1	-						380	-	-	0	0	457	0	-	1	-
108	-	0	0	0	176	-	-	1	-						381	-	-	0	0	458	1	-	1	-
109	-	1	1	1	178	0	0	0	-						383	0	0	-	-	459	1	-	1	-

Binary strings of the CDP for the maternally- and paternally-derived progeny chromosomes from complete, three-generation families. Grandmaternal origin is designated by 0, grandpaternal origin by 1; a dash indicates indeterminate grandparental origin. Abbreviations for the four loci are as in Table 1.

Table 3 Support for gene order

	log likelihood	Maximum likelihood recombination fraction (% recombination)		
BAIH	-53.4	B-A	A-I	I-H
		3.5	7.5	2.7
BAHI	-58.9	B-A	A-H	H-I
		3.5	8.2	2.6
BAIH/BAHI	5.5			

is 5.6, indicating the order B-A-I-H to be 400,000 times more likely than the alternative order. The same gene order is obtained by complete pedigree analysis with the pedigree analysis package (PAP)³² for the loci taken three at a time. The alternative order, B-A-H-I, would require five double exchanges to explain the observed pattern of recombinant chromosomes. It should be noted that the gene order B-H-I disagrees with previous work which suggests the order B-I-H^{28,29}.

To make these reference pedigree samples available to the scientific community, lymphoblastoid lines from many of the large families have been placed in the Human Genetic Mutant

Cell Repository, Camden, New Jersey. Furthermore, J. Dausset at the Centre for DNA Polymorphism, Hôpital St. Louis, Centre Hayem, 2 place du Dr Fournier, 75010 Paris, will make DNA from these families available as part of a large-scale collaborative effort to map the human genome.

We thank Debbie Meyers, John Edwards, Newton Morton,

David Botstein and Mark Skolnick for useful discussions, Mark Skolnick for help in sampling K1085, Mark Hoff for restriction enzyme preparations, and Joanna Rendi and Antonio Ferrey for help in preparing DNA samples. This work was supported by the Howard Hughes Medical Institute and a grant from the Institute of General Medicine, NIH.

1. Botstein, D., White, R., Skolnick, M. & Davis, R. W. *Am. J. hum. Genet.* **32**, 314-331 (1980).
2. D'Eustachio, P. & Ruddle, F. H. *Science* **220**, 919-924 (1983).
3. Harper, M. & Saunders, G. *Chromosoma* **83**, 431-439 (1981).
4. Bishop, D. T. & Skolnick, M. in *Banbury Report 4: Cancer Incidence in Defined Populations* (eds Cairns, J., Lyon, J. L. & Skolnick, M.) 421-433 (Cold Spring Harbor Laboratory, New York, 1980).
5. Lange, K. & Boenke, M. *Am. J. hum. Genet.* **34**, 842-845 (1982).
6. Bishop, D. T., Cannings, C., Skolnick, M. & Williamson, J. in *Statistical Analysis of DNA Sequence Data* (ed. Weir, B.) (Dekker, New York, in the press).
7. Murray, J. M. *et al. Nature* **300**, 69-71 (1982).
8. Kinston, H. M., Thomas, N. S., Pearson, P. L., Sarfarazi, M. & Harper, P. S. *J. med. Genet.* **20**, 255-258 (1983).
9. Gusella, J. F. *et al. Nature* **306**, 234-238 (1983).
10. Bhattacharya, S. S. *et al. Nature* **309**, 253-255 (1984).
11. Camerino, G., Mattei, M. G., Mattei, J. F., Jaye, M. & Mandel, J. L. *Nature* **306**, 701-704 (1983).
12. Bishop, D. T. & Gardner, E. J. in *Banbury Report 4: Cancer Incidence in Defined Populations* (eds Cairns, J., Lyon, J. L. & Skolnick, M.) 389-406 (Cold Spring Harbor Laboratory, New York, 1980).
13. Go, R. C. P., King, M. C., Bailey-Wilson, J., Elston, R. C. & Lynch, H. T. *J. natn. Cancer Inst.* **71**, 455-461 (1983).
14. Gottesman, I. & Shields, J. *Schizophrenia: The Epigenetic Puzzle* (Cambridge University Press, 1982).
15. Anderson, V. E. in *Genetic Basis of the Epilepsies* (eds Anderson, V. E., Hauser, W. A., Penry, J. K. & Sing, C. F.) 103-112 (Raven, New York, 1982).
16. Skolnick, M. H., Willard, H. F. & Menlove, L. A. *Cytogenet. Cell Genet.* **37**, 210-273 (1984).
17. Antonarakis, S. E., Boehm, C., Giardina, P. & Kazarian, H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 137-141 (1982).
18. Bell, G., Karam, J. & Rutter, W. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5759-5763 (1981).
19. Gerald, P. S. & Grzeschik, K. H. *Cytogenet. Cell Genet.* **37**, 103-126 (1984).
20. Goldfarb, M., Shimizu, K., Perucho, M. & Wigler, M. *Nature* **296**, 404-409 (1982).
21. Barker, D. & White, R. *Cytogenet. Cell Genet.* **37**, 412 (1984).
22. Ullrich, H. *Science* **209**, 612-615 (1980).
23. Barker, D., Holm, T., Jorde, L. & White, R. *Am. J. hum. Genet.* (in the press).
24. Morton, N. E. *Am. J. hum. Genet.* **7**, 277-318 (1955).
25. Thompson, E., Kravitz, K., Hill, J. & Skolnick, N. in *Genetic Epidemiology* (eds Morton, N. E. & Chung, C. S.) 247-253 (Academic, New York, 1978).
26. Edwards, J. H. *Ann. hum. Genet.* **34**, 229-250 (1971).
27. White, R. *et al. in Banbury Report 14: Recombinant DNA Applications to Human Disease* (eds Caskey, C. T. & White, R. L.) 235-250 (Cold Spring Harbor Laboratory, New York, 1983).
28. Fearon, E., Antonarakis, S., Meyers, D. & Levine, M. *Am. J. hum. Genet.* **36**, 329-337 (1984).
29. Gerhard, D., Kidd, J., Gusella, J. & Housman, D. *Proc. 16th Miami Winter Symp.*, 172-173 (1984).
30. Renwick, J. H. *Eur. Soc. hum. Genet.* **2**, 7-8 (1968).
31. Bridges, C. B. & Morgan, T. H. *The Third Chromosome Group of Mutant Characters of Drosophila melanogaster* (Carnegie Institute, Washington, 1923).
32. Hassstedt, S. J. & Cartwright, P. E. *PAP-Pedigree Analysis Package, Technical Report No. 13* (University of Utah, Utah, 1981).

ARTICLES

Optical dating of sediments

D. J. Huntley*, D. I. Godfrey-Smith† & M. L. W. Thewalt*

* Department of Physics and † Department of Archaeology, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6

A new method for dating sediments is proposed, which determines the time since the sediment was last exposed to sunlight. An argon-ion laser is used to excite electrons from thermally-stable light-sensitive traps and the subsequent luminescence used as a measure of the past radiation dose. Two sample sequences spanning the periods 0-700 kyr and 0-6 kyr show steadily increasing luminescence with age. An age of 62 ± 8 kyr is obtained for a silt radiocarbon dated at 58.8 ± 0.3 kyr. Problems were found with two much younger samples.

THE dating of geological and archaeological events would be easier if the mineral grains that form a sediment could be dated directly. At present, the dating of sediments relies almost entirely on radiocarbon dating of organic matter with its accompanying problems of scarcity, age limitation and association. We demonstrate here a new method for the direct dating of sediment grains. The zeroing mechanism is sunlight and the age obtained is presumed to coincide with the date of deposition of the sediment.

The new method falls in the same class of dating methods as thermoluminescence (TL) and electron spin resonance. The underlying principle is that the response of naturally-occurring minerals to environmental radiation is cumulative over time. Independent measurements can determine both the total radiation dose received by the mineral and the environmental dose rate it experienced *in situ*. The age can then be deduced¹. In detail, ionizing radiation resulting from the decay of the naturally occurring radioisotopes ⁴⁰K, ²³²Th and ²³⁸U interacts with a crystalline substance, freeing electrons from their normal atomic sites. Some of these electrons become trapped at defect sites and, if the trap depth is large enough, will remain there indefinitely on a geological time scale. The number of these trapped electrons is thus a measure of the radiation dose since some event in which all the traps were emptied.

In this new method, only those trapped electrons which can

be freed by absorption of visible photons through exposure to light are sampled. We refer to them as light-sensitive. The objective is to determine the radiation dose, and hence time, since the last exposure to sunlight. In this way, it should be possible to date a variety of archaeological and geological sediments and other objects which have been well covered or buried.

The essence of the method is to eject the trapped electrons using an intense light source and measure the luminescence resulting from their recombination. This physical principle has been used for several decades in infrared-visible image converters (see, for example, ref. 2).

In principle, the discrimination of induced luminescence from scattered incident light can be made on the basis of wavelength or time; we have decided to use the former.

Our initial experiments were designed to demonstrate the feasibility of this technique and were conducted with the apparatus shown in Fig. 1. A 514.5-nm beam from an argon-ion laser was passed through blue-absorbing filters to remove unwanted <500-nm light, and into the sample chamber where it was directed with a mirror onto a few milligrams of sample. The intensity at the sample was $\sim 50 \text{ mW cm}^{-2}$. The scattered 514.5-nm beam was attenuated by a factor of $\sim 10^{16}$ by blue-pass glass filters in front of the photomultiplier tube and yielded a count rate of $\approx 50 \text{ s}^{-1}$. The sample luminescence at $\sim 400 \text{ nm}$

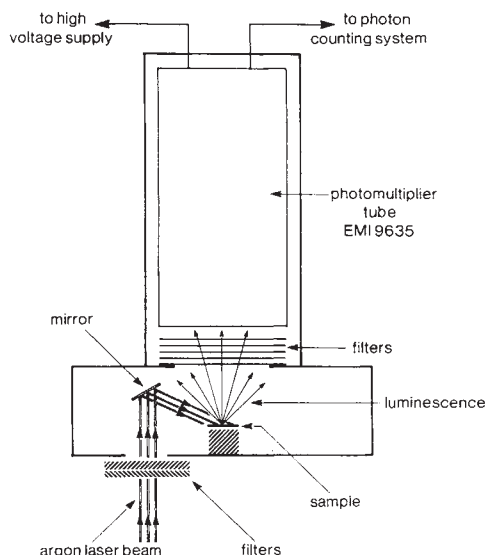


Fig. 1 Schematic of the apparatus. Various filter combinations were used. A 520-nm interference filter and Corning 3-70 and 3-71 glass filters were placed in the incident beam and a combination of Corning 7-59 and 5-58 filters were placed in front of the photomultiplier.

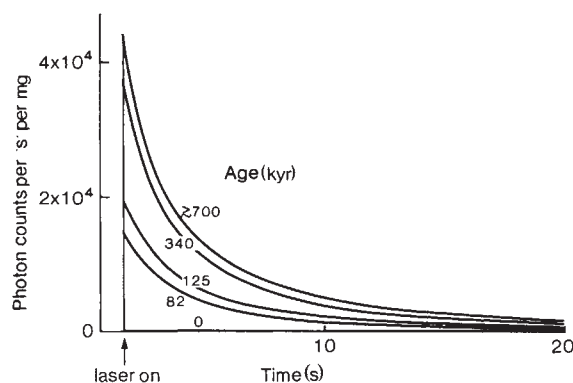


Fig. 2 Luminescence intensity versus time after the laser is switched on. The samples are 100- μ m quartz grains extracted from a sequence of stranded beach dunes in South Australia⁴ and show increasing intensity with age.

was attenuated by an order of magnitude by the same filters and yielded count rates up to 10^6 s^{-1} .

Two kinds of experiments were performed. In the first, we measured two sets of geological sediments, each of which contained several units representing discrete depositional events, to show the increase of light intensity with age. In the second, we used the method to determine a past radiation dose for three samples of known age and compared this with the expected dose calculated from the age, and to the dose obtained from previous TL measurements. Strictly, we determine the so-called 'equivalent dose' which is the laboratory γ dose that produces the same signal as the past environmental dose; the two differ because of different sample sensitivities to α , β and γ radiation.

Luminescence results

The first samples studied were from a sequence of stranded beach dunes in south-east South Australia. These were created at high sea-level stands during interglacial periods over the past million years and their ages may be deduced by matching the topography to the oxygen-isotope stages of Shackleton and Opydyke³ and to the magnetic boundary of Idnurm and Cook⁴. A detailed study of the younger dunes has been made by

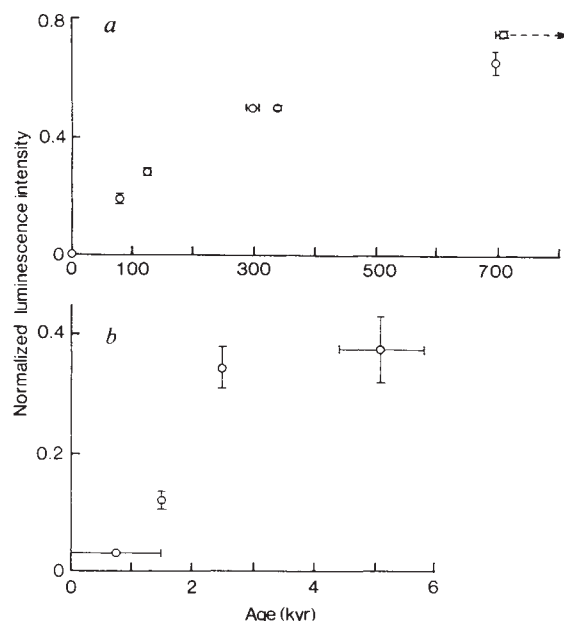


Fig. 3 Luminescence intensity versus age of quartz extracts for two sequences of geologically similar sediments: *a*, seven stranded beach dunes from south-east South Australia; *b*, four Ah soil horizons from the HaRk-1 site⁶ in British Columbia. Each sample's natural luminescence intensity is normalized to its luminescence response to a 30-Gy laboratory γ dose. This procedure was used to adjust for sensitivity variations between samples and reduced the scatter from duplicate samples. The luminescence increase with age observed for both sets of samples suggests that the phenomenon should be useful for dating on time scales of 10^3 – 10^5 yr.

Schwebel⁵. All the dunes are made of calcarenite from which pure α -quartz was extracted for this study.

Luminescence signals observed from some of these samples are shown in Fig. 2. After the laser beam is switched on there is an intense signal which decays non-exponentially. This type of signal is representative of all the samples studied here. The intensity of the signal as a function of age is shown in Fig. 3*a* and was found to increase smoothly from <30 photon counts per second (p.c.p.s.) per mg of modern beach sand to $\sim 4 \times 10^4$ p.c.p.s. mg^{-1} for the oldest one.

A similar study was made of quartz extracts from a sequence of Ah soil horizons formed on aeolian parent material in a stratified archaeological site⁶. The results are shown in Fig. 3*b*. The ages, ranging from 0 to 6 kyr, were deduced from ^{14}C ages from charcoal within or bracketing each horizon.

In the second type of experiment, we have attempted to use the method to determine the equivalent doses of three sediments and have compared them with those obtained by the R- Γ (or partial bleach) TL method^{7–9}. Equivalent doses were determined by preparing several similar specimens of each sample, administering to them a variety of ^{60}Co γ doses, heating them to 250 °C and then measuring the induced luminescence. The equivalent dose was obtained by extrapolating the luminescence intensity versus laboratory dose to zero intensity as shown in Fig. 4. The purpose of the heating was to empty electrons from traps that are not thermally stable on the relevant geological time scale. In this way, the luminescence should arise only from light-sensitive thermally-stable traps.

The first sample (KRGP-2) consisted of 2–8 μ m grains extracted from a layered pond silt immediately underlying a wood layer ^{14}C dated at 58.8 ± 0.3 kyr (QL-195). Divigalpitiya^{8,10} obtained an equivalent dose of 130 ± 22 Gy and an age of 64 ± 12 kyr for this sample using TL. Our results are shown in Figs 4 and 5*a*. The dose obtained was 125 ± 15 Gy, yielding an age of 62 ± 8 kyr in excellent agreement with the TL result and the ^{14}C age.

The final two samples selected were ones which could not be dated successfully by TL and we hoped that the new method

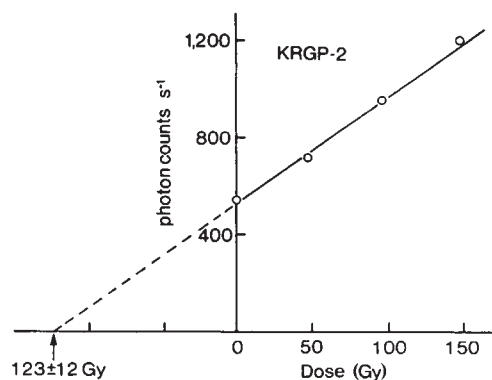


Fig. 4 Plot of luminescence intensity against laboratory ^{60}Co γ dose used to determine an equivalent dose for the KRGP-2 silt.

would give better results. The first of these, quartz grains extracted from the 5 kyr Ah soil horizon of Fig. 3b, gave an equivalent dose of 15 ± 1.5 Gy (Fig. 5b). This agrees with the TL result of 16 ± 2 Gy but is double that expected for its age (for dose-rate data, see sample HaRk-1D)⁸. We do not know whether this reflects a lack of understanding of quartz behaviour or whether it dates an earlier depositional event (even though the silt grains from the same horizon gave the correct age⁸).

The last sample was the quartz fraction from a present-day Pacific-coast beach sand (lat. $48^\circ 23' \text{N}$, long. $124^\circ 6' \text{W}$) and we, therefore, expected it to give an equivalent dose equal to zero. The value we obtained was 8 Gy initially, increasing with time to almost double this value after 12 s of laser illumination (Fig. 5c). The TL result was several times larger. We suggest that this sample has not been well-exposed to sunlight on the beach and that this fact can be deduced from the lack of constant equivalent dose with illumination time as traps progressively less light-sensitive are sampled.

Discussion

The inspiration for this research came from work in the development of thermoluminescence techniques for dating of sediments⁷⁻⁹, in which the R- Γ method was developed to extract the TL component arising from light-sensitive traps. Its usefulness is limited by the background TL signal from light-insensitive traps which may preclude extraction of the signal from the most light-sensitive traps. In contrast, in the new method the initial rapidly decaying signal arises from the most light-sensitive traps. Present indications are that this signal arises from traps more light sensitive than those accessible by TL.

As the signal decay shown in Fig. 2 is not a simple exponential, a range of trap sensitivities is indicated. The initial decay, however, arises from traps for which $\sim 10^{-1} \text{ J cm}^{-2}$ of 514-nm light empties $\sim 50\%$; for these we would expect more than 90% to be emptied by ~ 10 s of sunlight. Appropriate extraction of this signal should enable samples with depositional sunlight exposures as short as half a minute to be dated.

Tests for sufficient depositional light exposure and for thermal stability of the traps are possible with this method. The thermal stability test can be done by varying the pre-heat temperature, provided this does not cause significant transfer of electrons from light-insensitive to light-sensitive traps. If this occurs, optical draining of the thermally unstable traps may prove to be a better alternative.

Figure 2 shows that a short pulse of excitation yields a measurable signal without significantly depopulating the traps. This gives rise to the possibility that the signals from both natural and laboratory radiations can be measured using the same sample disk; this would be particularly valuable for very small samples or those of highly variable sensitivity.

This new method can, in principle, be used for dating other events such as heating and growth for which TL is presently used, and offers several advantages. There is no interfering incandescence; problems with thermal contact and sample alter-

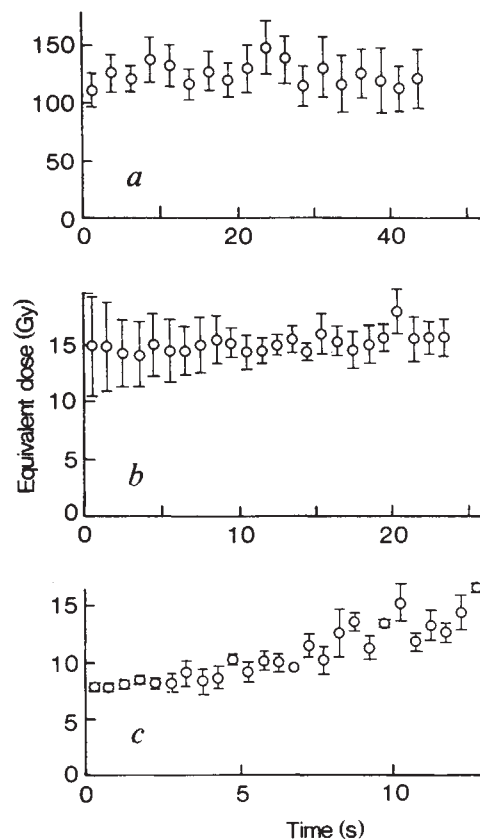


Fig. 5 Plots of equivalent dose against time after the laser is switched on for three samples. a, KRGP-2, a 59-kyr silt; b, quartz from HaRk-1 D, a 5-kyr Ah soil horizon; c, quartz from CBBS, a present-day beach sand.

ation and/or oxidation during heating will be reduced or eliminated. Automated measurements will be possible for all types of samples. A disadvantage will be the fact that, as only light-sensitive traps are sampled, the method will be more prone to error due to accidental light exposure.

In conclusion, the optically-induced luminescence of naturally occurring minerals is readily measurable and its intensity increases with age for sediments on a time scale of 10^3 – 10^5 yr. The method should be well-suited to dating sediments which have received at least a short exposure to sunlight during the depositional event, and may also be applicable to samples for which thermoluminescence or electron spin resonance is presently used. The simplicity of the measurement lends itself readily to rapid and automated measurements. The pilot work described here has demonstrated the feasibility of the method and the validity of the underlying principles. However, much further detailed study is required to gain a thorough understanding of the behaviour of the minerals before routine dating can be attempted.

We thank W. M. R. Divigalpitiya for use of his samples; D. Robert and L. Veilleux for assistance; D. Alexander, J. E. Armstrong, S. Hicock, J. Hutton and J. R. Prescott for assistance with sample collection. D.J.H. thanks J. R. Prescott and the University of Adelaide for hospitality and assistance in sample preparation. The research was supported by the Natural Sciences and Engineering Research Council of Canada.

Received 17 September; accepted 22 October 1984.

1. Aitken, M. J. *Physics and Archaeology* 2nd edn (Oxford University Press, London, 1974).
2. Leverenz, H. W. *An Introduction to Luminescence of Solids* (Wiley, New York, 1950).
3. Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* 3, 39–55 (1973).
4. Idnurm, M. & Cook, P. J. *Nature* 286, 699–702 (1980).
5. Schwebel, D. A. thesis, Flinders Univ. of South Australia (1978).
6. Valentine, K. W. G., Fladmark, K. R. & Spurling, B. E. *Can. J. Soil Sci.* 60, 185–197 (1980).
7. Wintle, A. G. & Huntley, D. J. *Can. J. Earth Sci.* 17, 348–360 (1980).
8. Huntley, D. J., Berger, G. W., Divigalpitiya, W. M. R. & Brown, T. A. *PACT* 9, 607–618 (1983).
9. Wintle, A. G. & Huntley, D. J. *Quat. Sci. Rev.* 1, 31–53 (1982).
10. Divigalpitiya, W. M. R. thesis, Simon Fraser Univ. (1983).

Genetic organization of *Drosophila* bithorax complex

E. Sánchez-Herrero*, I. Vernós†, R. Marco† & G. Morata*

* Centro de Biología Molecular del CSIC, Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain

† Instituto de Investigaciones Biomédicas del CSIC y Departamento de Bioquímica de la Facultad de Medicina, Universidad Autónoma de Madrid, 28029 Madrid, Spain

The Drosophila bithorax complex is subdivided into three major genes: Ultrabithorax⁺, abdominal-A⁺ and Abdominal-B⁺. Each of these genes plays its principal part in a particular anatomical domain of the body, where it is specifically required to determine the correct segmental pattern.

GENE control of morphogenesis is one of the fundamental problems of developmental biology. The bithorax complex (BX-C) of *Drosophila*¹⁻³ is a group of related homoeotic genes determining the morphological pattern of two thoracic (T2 and T3) and eight abdominal (A1-A8) segments. Previous work¹⁻³ has shown that segment diversity and identity, but not segmentation itself, depends on the normal function of the BX-C. Several mutations have been isolated which inactivate discrete functions of the BX-C as their phenotypes are subsets of that of the entire deletion of the complex. For example^{1,2}, mutations at the *bx* locus transform T3a into T2a and *pbx* mutations transform T3p into T2p (a, anterior and p, posterior compartment). Thus, the loss of single BX-C functions only affects one or two of the compartments or segments controlled by the complex as a whole. There are, however, other mutations that appear to inactivate simultaneously several BX-C functions. *Ultrabithorax* (*Ubx*) mutations, for example, inactivate *bx⁺*, *pbx⁺* and *bx^d⁺* functions^{1,2}, suggesting that there exist hierarchies of genetic elements within the BX-C.

The entire deletion of the BX-C results in a homoeotic transformation of all the segments (T2 to A8), indicating that the BX-C is responsible for their morphological diversity. Although there are some individual BX-C mutations affecting abdominal segments^{2,4,5}, evidence for these is incomplete. The number of abdominal functions encoded in the BX-C and their organization are unknown. *Ubx* mutations complement the BX-C mutations affecting abdominal segments⁵ and the BX-C can be split into two autonomous domains⁶; this suggests a functional independence of BX-C thoracic and abdominal genes.

To investigate these suggestions, the complete genetic and developmental description of all the functions of the BX-C is required. In particular, because defects in major functions, for example *Ubx*, are lethal, we seek to elucidate how many of these lethal complementation groups there are in the BX-C and what effects these mutations have on segmental development.

We find that the BX-C can be subdivided into three complementation groups: *Ubx*, *abdominal-A* (*abd-A*) and *Abdominal-B* (*Abd-B*). The latter two groups specify the development of the abdominal segments and at least part of the genitalia. Each of the three genes may contain a set of subordinate functions determining the development of individual compartments or segments.

DfP9 complementation groups

All the mutations reported here are induced on a genetically homogeneous chromosome carrying several markers⁷. We have used ethyl methane sulphonate (EMS) and X rays as mutagens and have detected the mutations in a F₂ screen as either lethal or viable (see Fig. 1 for experimental details). From 12,824 mutagenized chromosomes tested (7,280 mutated by EMS, 4,264 by X-ray), 41 mutations were isolated and located subsequently in three intervals within *DfP9* defined by the breakpoints of *DfP9*, *Dfbsd*¹⁰⁰, *Df109* and *DfC4*. Once a group of mutants was located on an interval, their complementation pattern was studied. Apart from two large deficiencies which could not be

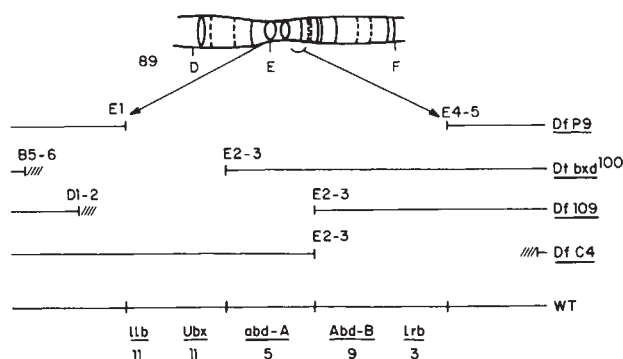


Fig. 1 Deficiency mapping and complementation analysis of the new BX-C mutants. To ensure genetic homogeneity, all the new mutants described here were induced on a multiple marked (*mm*) chromosome carrying the markers *mwh jv st red sbd² e¹¹ ro ca* (see ref. 13 for their description). This stock was built up from a single recombinant male. We used a regular F₂ screen to isolate lethal or visible mutants on a chromosome fragment; *mm* males were treated with either EMS (0.025%) or X rays (3,500–4,000 R) and crossed to females with marked third chromosomes such as *TM1* or *Mc* (*Microcephalus*). *Mc* alters head structures and covers the dominant sterility of deletions of the distal elements of the BX-C; it was used to ensure that we recovered any lethal or visible dominant sterile BX-C mutant. F₁ males, *mm/TM1* or *mm/Mc*, were then crossed individually to females carrying *Dp(3;3)P5/Df(3R)P115* where *Df(3R)P115* is a deletion extending from 89B7-8 to 89E7-8 and therefore includes the entire BX-C and several other neighbouring genes. In those cases where in the F₂ the class of flies *mm/Df(3R)P115* was either missing or showed mutant traits, *mm/Dp(3;3)P5* siblings were selected (*Dp(3;3)P5* is a tandem duplication of the BX-C; ref. 15) and a balanced stock was established. The mutant chromosomes were tested subsequently against *Df(3R)P9*, *Df(3R)bsd¹⁰⁰*, *Df(3R)109* and *Df(3R)C4* (described in refs 2, 3, 8, 15; designated in the figure and throughout the text in the abbreviated form) and located to the intervals defined by these breakpoints. The mutants in each interval were then crossed *inter se* to determine the number of complementation groups per interval. In the intervals containing more than one complementation group, the genetic order cannot be established by complementation analysis. In the case of *llb* and *Ubx* we know that *llb* is to the left, because *llb* is viable over *Ubx^{C1}* which is a deletion of part of the *Ubx* and the *abd-A* genes, this latter undisputably to the right of *Ubx* as shown by deficiency mapping. The location of *lrb* to the right of *Abd-B* is based on the observation that in *Df(3R)M1* (J. Casanova and G.M., unpublished data) *Ubx⁺*, *abd-A⁺* and *Abd-B⁺* are deleted but *lrb⁺* is present.

assigned to a particular group, the remaining 39 mutations were classified as belonging to one of five complementation groups (Fig. 1). We name the genes corresponding to these five groups as follows: *lethal left of bithorax* (*llb*), *Ubx*, *abd-A*, *Abd-B* and *lethal right of bithorax* (*lrb*). As we have isolated several mutant alleles for each locus, we are confident that we have found all loci in *DfP9* causing lethality or abnormal pattern, unless their

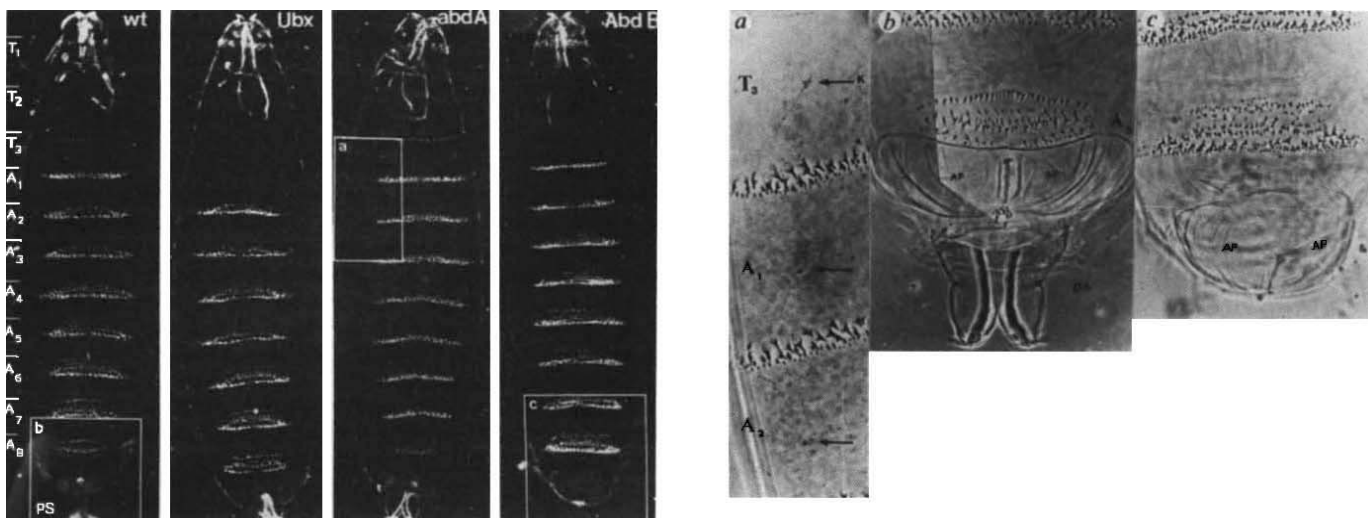


Fig. 2 Ventral cuticular patterns of wild type, *Ubx*⁻, *abd-A*⁻ and *Abd-B*⁻ larvae. The *Ubx*⁻ and *abd-A*⁻ larvae were taken out of the egg case before fixation. All larvae were mounted as described previously²⁵. wt, Wild-type first-instar larva. For a detailed description of the cuticular pattern, see ref. 26. A, abdominal segment; T, thoracic segment ($\times 60$), *Ubx*, ventral cuticle of a hemizygous *Ubx*¹³⁰/*DfP9* larva. The transformation of T3 and A1 denticle belts toward mesothorax is visible, as well as the presence of Keilin's organs in the A1 segment. This organ has only two hairs, indicating that the transformation is restricted to the anterior compartment. Ventral pits are visible regularly in the abdominal segments of these transformed larvae. The abdominal segments posterior to A1a also show some degree of anterior transformation particularly on the dorsal side, but this effect is visible also in *DfP9*/+ larvae and is thus attributable to a dominant effect of the deficiency (I.V., and R.M., unpublished results). *abd-A*, Ventral cuticle of an *abd-A*^{M1}/*Df109* larva. The same phenotype is visible in *abd-A*^{M1}/*abd-A*^{MX1} larvae. A clear transformation of the denticle belts of A2–A7 towards A1 can be seen. Segment A8 denticle belt is also transformed but it is not clear to which wild-type segment. **a**, Detail of the metathoracic and first two abdominal segments of the *abd-A*⁻/*Df109* larva: the Keilin's organ in T3 and the corresponding monohairs in A1 and A2 are indicated by arrows. These monohairs reveal a transformation of the posterior compartments of the abdominal segments towards posterior metathorax. These monohairs appear with some degree of variability in any segment A2–A7, indicating that all these segments are transformed, albeit in part, to a posterior thoracic segment, probably T3p ($\times 240$). *Abd-B*, Ventral cuticle of *Abd-B*^{M1}/*Abd-B*^{M5}. The same transformation is visible in segments A4–A8 of *Abd-B*^{M5}/*DfP115* larvae. Their denticle belts are almost identical and do not show the progressive reduction in size of the more posterior rows of denticles to the same degree as that seen in the wild-type larvae. **b**, Detail of the eighth abdominal segment of a wild-type *Drosophila* larva ($\times 160$). **c**, Eighth abdominal segment of a *Abd-B*⁻ homozygous larva showing the naked cuticle behind the eighth abdominal denticle belt, the normal anal pads (AP) and the absence of posterior spiracles (ps) and corresponding filzkörper ($\times 160$).

mutation rate is much lower than those of the loci described here. Loci affecting other functions, such as behaviour, would not have been detected.

To define the effect of the mutations on the larval segments, we next studied the cuticular pattern of late embryos (after secretion of the larval cuticle) or of first-instar larvae. Like *DfP9* (refs 2, 8), all the mutants were expected to reach late embryonic stage. As *DfP9* embryos secrete cuticle and are composed of a normal number of segments, we expect none of the mutations to cause cuticular defects or to alter the number of segments.

We studied the effect on adult segments by generating mutant clones in flies of normal phenotype which carry a normal dose of the BX-C⁺ (Table 1, Fig. 3). It has also been studied in complete flies when the mutations were viable or produced a small percentage of surviving flies (escapers).

We began with the two complementation groups that seemed to map at opposite ends of *DfP9*. Ten mutations (nine induced by EMS, one by X ray) of the *llb* gene have been isolated. The embryonic phenotype of two alleles of *llb* over *DfP9* was normal, indistinguishable from wild type. Three mutant alleles (two EMS, one X ray) of the *lrb* gene have been isolated. As with *llb*, the embryonic phenotype was normal. Neither *llb* nor *lrb* seemed to be functionally related to BX-C as they did not affect segment development. We cannot, however, rule out the possibility of an effect on other structures, such as the terminalia, which cannot be scored easily in larvae.

BX-C complementation groups

We have isolated 11 lethal or viable mutations within the interval *DfP9*–*Dfbxd*¹⁰⁰ of *Ultrabithorax* that also failed to complement other known *Ubx* mutations. The *Ubx* lethal mutations have been described previously^{1–3,6,7,9–12}. Compartments T1a, T1p and T2a remained unaffected, T2p was transformed into T1p,

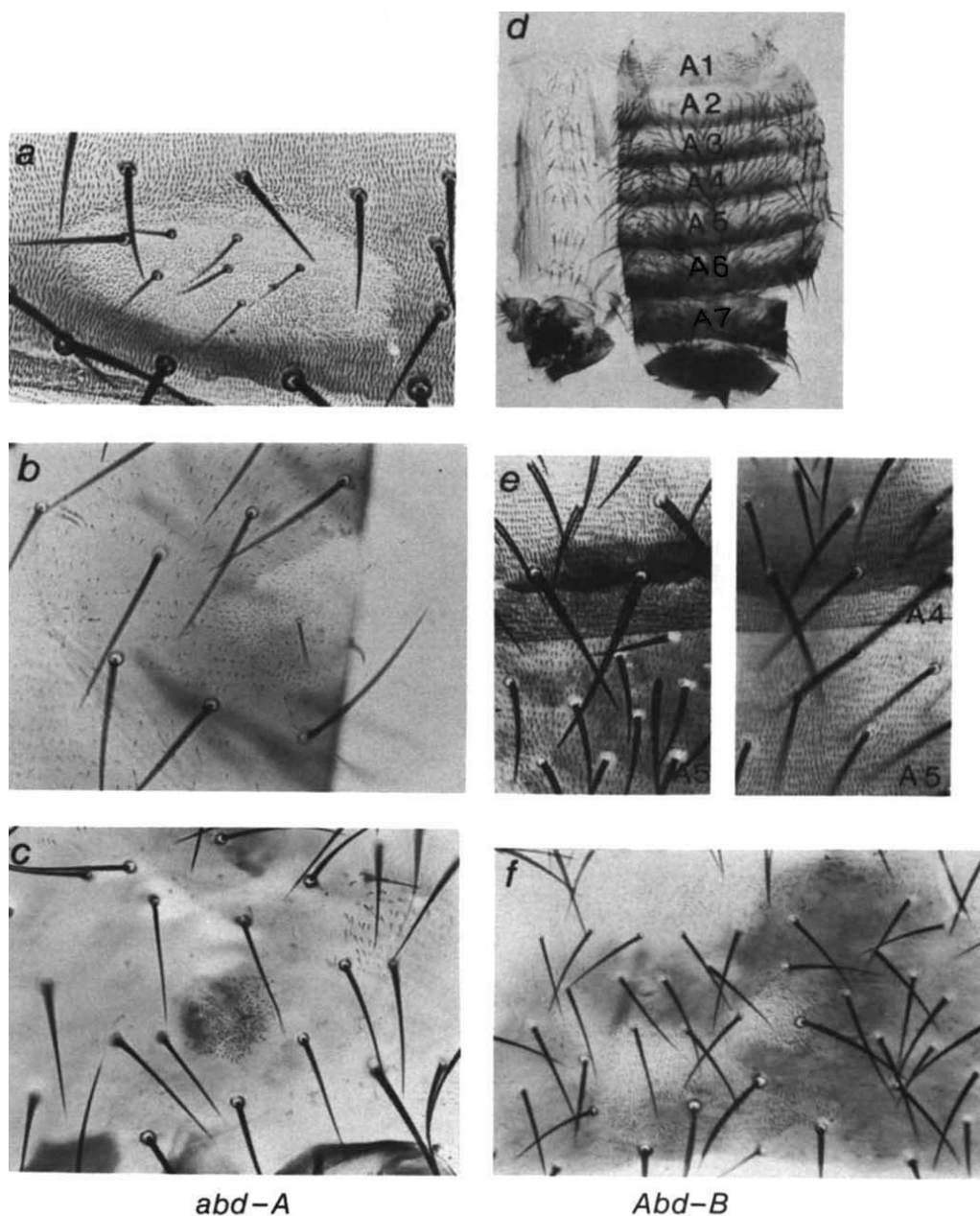
T3a into T2a, T3p into T1p and A1a into T2a (Fig. 4). This syndrome is the sum of the four independent transformations of the BX-C thoracic genes *abx*, *bx*, *pbx* and *bxd*. The five new X-ray-induced *Ubx* lethals (*Ubx*^{MX1}, *Ubx*^{MX2}, ..., *Ubx*^{MX5}) fit this description, as they failed to complement individual *abx*, *bx*, *bxd* and *pbx* mutations and were lethal over *Ubx*¹.

The phenotype of *Ubx*⁻ embryos paralleled that of the adult, although in having fewer landmarks in the larval cuticle, the transformation of posterior compartments was more difficult to assess. Our findings confirm previous reports on the phenotype of *Ubx*⁻ embryos: the prothoracic transformation of T2p and T3p (ref. 12), the restriction of its homoeotic effect in A1 to the anterior compartment^{6,11,12} and a slight effect on the remaining abdominal segments^{2,6}.

We have found also six viable mutations in this complementation group (*Ubx*^{M1}, *Ubx*^{M2}, ..., *Ubx*^{M6}). They were all EMS-induced and did not behave as alleles of the classical *bx*, *pbx* or *bxd* loci; they showed the quadruple combination of *abx*, *bx*, *pbx* and *bxd* phenotypes in a mild form. Over individual mutations of the *bx*, *pbx* and *bxd* type, they produced a slight phenotype of the mutant in question. In addition, *Dpbxd*¹⁰⁰, carrying a proximal fragment of the BX-C¹³ containing the DNA corresponding to *abx*⁺, *bx*⁺ and *Ubx*⁺, covered the *abx* and *bx* components of the phenotype but not the *bxd* and *pbx* ones. All known *Ubx* mutants inactivated the *bxd* and *pbx* functions^{1,2,7} which in some cases were located tens of kilobases away from the mutated *Ubx* site¹⁴. All these observations indicated that the visible EMS-induced mutations were weak but viable *Ubx* alleles.

In this screening we have not recovered mutations affecting individual functions of the *abx*, *bx*, *pbx* or *bxd* type. Many of these individual mutations were produced by insertions¹⁴ and it seems unlikely that EMS or X rays could induce such lesions

Fig. 3 Homoeotic effect of *abd-A* and *Abd-B* mutants in the adult abdominal segments. *a-c*, *abd-A* mutants; *d-f*, *Abd-B* mutants. In the abdominal segments of the adult, the posterior compartments have been described only in the case of A1 (ref. 27), where it is a narrow band of bristle-free unpigmented cuticle. In the remaining segments it is not yet possible to determine the position of the anteroposterior boundary. As the posterior compartments are morphologically similar, a transformation of, say, A3p into A1p would not be discerned. Therefore, the clones in A2-A8 can only detect anterior transformations. *a*, Clone of genotype *abd-A^{M1}/DfP115*, marked with γ , *mwh* and *ju* (see Table 1 legend and ref. 5). The clone appears in A4 but differentiates as in A1, judging from the bristle size and pigmentation. Note that a wild-type bristle (γ^+), has penetrated the mutant clone, but retains its segmental character. *b*, Clone of the same genotype as *a* but generated in A5 of a male, suggesting that depigmentation is not completely cell autonomous. Although in the anterior part of the clone the line of depigmentation coincides sharply with the mutant trichomes, in the posterior side several trichomes appear pigmented. *c*, Typical *abd-A⁻* clone of the same genotype as *a* and *b* in segment A6. Clones formed in this segment produce trichomes in areas which in A6 are devoid of them, thus are indicative of a transformation towards a more anterior segment. They are of small size, generally rounded and sometimes produce 1-2 bristles (see ref. 5) larger than A1 bristles. The morphological features of these clones suggest a transformation towards A2-A5 associated with poor viability and abnormal differentiation. This may result in the consistently lower yield of clones in A6 than in more anterior segments (Table 1). *d*, Abdomen, open in dorsal and ventral halves, of a male of genotype *Abd-B^{M3}/DfP115*. Note that in addition to the terminalia (TER), this male has eight abdominal segments (wild-type males have only six, since the primordia for A7 and A8 do not develop). In females, however, the genitalia are usually absent or reduced, although analia are normal. This is consistent with the current view of the organization of terminalia which considers female and male genital primordia as separate, deriving from A8 and A9 primordia, respectively^{28,29}. Although the pattern of segments A3 and A4 is indistinguishable, we assume that the transformation is towards A4, because the A3 pattern requires *abd-A⁺* but not *Abd-B⁺* function. (The *abd-A⁻* mutants affect A3 pattern whereas *Abd-B⁻* mutants do not.) If A3 is included in the domain of *abd-A*, the identical A4 pattern should also be included. *e*, Comparison of the trichome density in A4 and A5 of left, a wild-type male and right, an *Abd-B^{M3}/DfP115* male. The trichomes in A5 are more spaced than in A4 for normal males. The difference is eliminated in the mutant flies with the transformation. *f*, Clone of genotype *Abd-B^{M1}/DfP115* in A6. In the area where the clone is situated, A6 do not contain trichomes but they appear in the clone, thus indicating a transformation towards a more anterior segment. To assess the pigmentation of the mutant clones (to decide whether a transformation has been towards A4 or A5), the mutant clones in this experiment have been marked with *mwh* and *ju* but not with γ , which alters pigmentation. The clone shows clearly depigmented areas, indicating that the transformation is towards A4.



on the DNA. However, some *bxd*, *pbx* and *abx* alleles have been X-ray-induced^{3,15}. Our results may reflect simply the different probability of mutation. *Ubx* mutations are two orders of magnitude more frequent than *bx* and *pbx* ones¹⁶. Moreover, the size of the DNA target for *Ubx* is larger than that of *abx* or *bxd* mutations¹⁴.

We have isolated five alleles of the *abd-A* gene (three EMS-induced: *abd-A^{M1,M2,M3}* and two X-ray-induced: *abd-A^{MX1,MX2}*). We have found also that the breakpoint of *T(2;3)P10* behaved as a lethal allele of this locus. The larval and adult phenotypes of this breakpoint have been described⁵.

The larval phenotype is defined on the basis of *trans*-heterozygotes of the lethal alleles *abd-A^{M1}* and *abd-A^{MX1}* and their respective hemizygous phenotype over *Df109* (Fig. 2). These combinations gave a phenotype indistinguishable from that of

DpP10;Df109, a deletion for at least part of the *abd-A* DNA⁵. This suggested that the two mutants studied are null alleles of this gene and that their phenotype corresponded to the complete loss of function of the *abd-A* gene. In the larva all the head and thoracic segments appeared normal. The abdominal segments were affected and displayed a similar pattern: A1a in the anterior region and T3p in the posterior, both dorsally and ventrally (Fig. 2). The transformation was less complete posterior to A4, particularly on the dorsal side. All structures posterior to A8 appeared normal.

The effect of loss of *abd-A⁺* on the adult cuticle has been studied by genetic mosaics. Mutant clones of *abd-A^{M1}* were generated during the larval period. The results paralleled those found for the larval segments (Table 1). Clones in head, thorax, A1a and genitalia differentiated normally. Clones in A2-A4

Table 1 Number of *abd-A*⁺ and *Abd-B*⁺ clones showing homoeotic transformations/total number of clones per segment

Genotype	n	H	T2	A1	A2	A3	A4	A5	A6	A7	TER
<i>abd-A</i> ⁺ /DfP115	112	0/11	0/27	0/15	32/32	35/35	32/32	22/22	6/6	—	0/3
<i>Abd-B</i> ⁺ /DfP115*	260	0/9	0/74	0/20	0/21	0/20	0/27	11/12	31/32	15/15	0/4
<i>DpP115/+; abd-A/DfP115</i> (control)	80	0/6	0/28	0/11	0/23	0/18	0/24	0/11	0/20	0/10	0/3

n, Number of flies examined; H, head-antennal segment; T2, mesothoracic segment; A1–A8, abdominal segments; TER, analia and genitalia. *abd-A*⁺ clones were generated by mitotic recombination, induced by X-ray (1,000 R) in male larvae (48–120 h after egg laying) of genotype *y; Dp(1;3)sc¹⁴Dp(3;3)146M(3) i⁵⁵ Df(3R)P115/mwh ju st red sbd² abd-A^{M1}e¹¹ ro ca*. This genotype is viable and has a normal segmental pattern. A virtually identical scheme of mitotic recombination has been described recently in detail⁵. The production of marked clones rests on the elimination by mitotic recombination of the *Dp(3;3)146* that contains a full dose of the BX-C⁺ and therefore covers the lethality of the combination *abd-A^{M1}/Df(3R)P115*. The marker mutants are arranged in such a manner that the *abd-A*⁺ clones are characteristically labelled and at the same time are given a growth advantage over the rest of the cells²⁴. *Abd-B*⁺ clones are generated by the same procedure in male and female larvae, but replacing *abd-A^{M1}* by *Abd-B^{M1}*. In the genotype of controls, an extra dose of the BX-C⁺ (*Dp(3;1)P115*) was introduced into female flies so that the marked clones still contain the BX-C⁺ function, verifying that any homoeotic effect shown by the clones is due to the mutation covered by the BX-C⁺ functions.

* Only clones on the dorsal side (tergites) have been included. The number of clones in H and T2 is based on a smaller number (115) of flies.

differentiated structures typical of A1 both dorsally and ventrally (Fig. 3, Table 1). In segment A5 the mutant clones frequently produced an incomplete transformation towards A1 showing, for example, thin short bristles typical of A1 together with other longer ones characteristic of more posterior segments. In segment A6 they formed a pattern unlike A1 but similar to any segment posterior to A1 and anterior to A5.

As the allele *abd-A^{M3}* is hemizygous viable, the phenotype of adult flies can be studied. The homoeotic transformation was similar but much weaker than that seen in the clones of *abd-A^{M1}/Df115*. The mutant flies show pattern elements characteristic of A1 in segments A2–A6; dorsally, bristles of small size similar to those of A1 frequently appeared and there was a marked reduction in the number of ventral bristles in the sternites. For example, in males there was a mean of 3.6 ± 1.3 bristles (wild type = 11.6 ± 1.1 bristles) in A2 and in the A6 of females there were 5.1 ± 2.0 bristles (wild type = 22.6 ± 1.6). This result suggested that sternites of segments A2–A6 were partially transformed toward the sternite of A1 which was devoid of bristles.

We have also studied the phenotype of flies *infraabdominal-2* of *Kuhn(iab-2^k)* in trans with *abd-A^{M1}*, *abd-A^{M2}* and *Df109*. All these combinations showed essentially identical phenotypes consisting of a partial transformation of A2 towards A1 as expected⁴. No effect was seen in any other abdominal segment posterior to A2. The lack of effect on segments other than A2 did not result from partial expressivity of *iab-2^k* as the transformation in A2 of *iab-2^k/abd-A^{M1}* was stronger than that of *abd-A^{M3}/Df109*, indicated by the greater reduction in bristles per sternite in A2 of females (wild type = 15.5 ± 1.0 , *iab-2^k/abd-A^{M1}* = 0.7 ± 0.8 and *abd-A^{M3}/Df109* = 3.8 ± 1.7). Yet *iab-2^k* is restricted to A2, whereas *Abd-A^{M3}* is not (in A3, wild type = 22.3 ± 2.0 , *iab-2^k/abd-A^{M1}* = 19.3 ± 2.4 and *abd-A^{M3}/Df109* = 8.5 ± 2.8 bristles per sternite in females). These results indicated that *iab-2^k* is included within the *abd-A⁺* gene but its phenotype is only a part of that of an *abd-A* mutation. This, in turn, suggested that *abd-A⁺*, like *Ubx⁺*, contained a set of distinct individual functions, controlling individual segments or compartments.

We have isolated nine alleles of the *Abd-B* gene, seven induced by EMS (*Abd-B^{M1,M2,...,M7}*) and two by X rays (*Abd-B^{MX1,MX2}*). All mutations showed a dominant trait also present in heterozygous deficiencies such as *DfC4* or *DfP9*, indicating that the locus is haplo-insufficient. Some of the lethal mutations produced escapers, so the alteration of the adult segment pattern could be studied directly. The adult phenotype of the trans-combinations over *DfP9* was a transformation of A5–A8 towards A4 (Fig. 3). However, the homoeotic effect of the trans-heterozygotes is not entirely homogeneous. Some combinations appeared to produce a phenotype intermediate between A4 and A5, judging from male pigmentation and trichome density. Transformation towards A5 has also been observed in the muscles¹⁷; there is a male-specific muscle in A5 reiterated in A6 and A7 in *Abd-B^{M5}* homozygous tissue (but not in *Abd-B^{M5}/DfP9* where

the male-specific muscle was absent or reduced (P. Lawrence, personal communication) in accordance with the transformation in the cuticle). The mutation *iab-5* transformed A5, and to a lesser extent A6 and A7, towards A4 (ref. 3). Similarly, our mutation *Abd-B^{MX2}* affected A5–A7 but, unlike other *Abd-B* alleles, did not affect female genitalia. We believe that these viable mutations are all weak alleles of *Abd-B*, although it is also possible that some of them affect discrete functions within *Abd-B* gene.

The dominant phenotype of *Abd-B* mutations is a slight transformation of the same type as that described for the escapers: in females, tergite of A6 and A7 differentiated some pattern elements characteristic of tergite A4 or A5. Males presented bristles in sternite A6 (normally devoid of them) and also developed a rudimentary tergite in A7 that showed features (pigmentation) of tergite A5 or A6.

It is possible, however, that the phenotype of escapers was not truly representative of the lack of function of the *Abd-B* gene, as escapers were more likely to appear from weak alleles. Therefore, we have studied the phenotype in genetic mosaics of a strong allele, *Abd-B^{M1}*, which does not produce escapers over the deletion of the gene. Clones hemizygous for this allele differentiated normally in head, thorax and abdominal segments A1–A4 (Table 1). In segment A5 they were almost always depigmented, indicating a transformation towards A4. Those in A6 and A7 were always transformed into A4 or A5, as they differentiated trichomes in areas where A6 or A7 lack them. As many clones in males were clearly depigmented (Fig. 3f), we conclude that the loss of *Abd-B⁺* results in a transformation towards A4.

The segmental transformations observed in the larval cuticle are consistent with those seen in the adult (Fig. 2). All the abdominal segments posterior to A5 showed a transformation towards a more anterior pattern similar to that of A4 or A5. The exact identity was not determined as the intermediate abdominal segments were almost indistinguishable, but there was no doubt that A7 and A8 were transformed. Posterior to A8 only the spiracles were affected, being very reduced or eliminated, consistent with the idea that the spiracles correspond to the posterior part of the eighth abdominal segment¹⁸.

Major BX-C genes and domains

Our results indicate that the BX-C can be subdivided into three independent genetic units: *Ubx⁺*, *abd-A⁺* and *Abd-B⁺*. The mutations in any of these genes affect several segments, yet they are most probably due to a single event. This is in contrast with current models of the BX-C function^{2,3,6,12} which assign one gene per segment or per segmental unit. It cannot be argued that our mutations are deletions for more than one gene as we have recovered weak alleles for each of the three genes, showing minor but clear effects extending over several segments. This was particularly clear in the case of the viable *Ubx* mutants and in *abd-A^{M3}*. In both cases, deletions of the genes occurred

(*DfP10* for *Ubx*, *DpP10*; *Df109* for *abd-A*) and the phenotype of the deletions was seen in the same segments but was much stronger than that of viable mutants.

Our analysis of the mutant phenotypes indicates that each of the three genes is principally responsible for the normal pattern of a particular region of the fly, thus defining three anatomical domains in which the entire realm of function of BX-C can be subdivided.

The *Ubx* domain (Fig. 4) extends from the anteroposterior boundary of T2 to the anteroposterior boundary of A1 (refs 6, 7, 11, 12). The elimination of *Ubx*⁺ function results in a series of four compartments T2p-T3a-T3p-Ala developing as T1p-T2a-T1p-T2a, defining a ground pattern of T1p for posterior and T2a for anterior compartments. Anterior to T2p, *Ubx*⁺ plays no role in segment development^{2,7}, even on structures morphologically like those of the domain when produced in a more anterior position¹⁹. The elimination of *abd-A*⁺ and *Abd-B*⁺ had no effect on the *Ubx* domain, indicating that *Ubx*⁺ is the only BX-C gene with a functional role in this region.

The *abd-A* domain extends from A1p-A4. In this region, the elimination of the *abd-A*⁺ gene caused the anterior compartments to develop as A1a and the posterior compartments as T3p. This ground pattern corresponded to that of *Ubx*⁺, indicating that although the *Ubx*⁺ product is present in the *abd-A* domain, the specific pattern of A1a-A4 segments was conferred by the function of *abd-A*⁺. When both *Ubx*⁺ and *abd-A*⁺ genes were absent as in *Df109* (refs 2, 5, 8), the compartments of the two domains developed equally as a T1p-T2a series.

However, the role of *Ubx*⁺ in the *abd-A* domain is minor; the combination *Ubx*⁻ *abd-A*⁺ in the *abd-A* domain showed a pattern indistinguishable in adult segments from that of *Ubx*⁺ *abd-A*⁺ and only a slight though significant thoracic transformation in the larval segments^{2,6}. This contrasts with the effect observed in the combination *Ubx*⁺ *abd-A*⁻ where the whole segment pattern was profoundly transformed (Figs 2, 3). The *Abd-B* mutations had no effect on the *abd-A* domain. We conclude, therefore, that the pattern of this domain is determined by the *abd-A*⁺ gene.

Finally, the *Abd-B* domain extends from A5 to A8, defined by the realm of action of the *Abd-B*⁺ gene. The elimination of the *Abd-B*⁺ product resulted in all the segments of the domain developing as A4, a pattern specified by *Ubx*⁺ and *abd-A*⁺. This indicates that, although these products are present in the *Abd-B* domain, the specific identity of segments A5-A8 depends on the *Abd-B*⁺ function.

It is difficult to assess the contribution of *Ubx*⁺ and *abd-A*⁺ genes to the specification of the *Abd-B* domain. We believe the role of *Ubx*⁺ to be minor. However, the *abd-A* mutations had a clear effect on segments A5-A8, so the *Abd-B*⁺ gene product is necessary but not sufficient for the correct segment pattern in the *Abd-B* domain. *abd-A* mutations may inactivate partly the adjacent *Abd-B*⁺ gene by the phenomenon of *cis*-vection, which has been described for some mutations of the BX-C²⁰. Thus, part of the effect exerted by the *abd-A* mutant on the *Abd-B* domain could be attributed to a partial inactivation of the *Abd-B*⁺ gene. On the other hand, *abd-A*⁺ (and *Ubx*⁺) may provide in the *Abd-B* domain the ground pattern necessary for the *Abd-B*⁺ gene to act and confer the specificity. This would be in agreement with a combinatorial model on BX-C function^{6,21}.

It is worth noting that the sum of the mutant phenotypes of *Ubx*⁻, *abd-A*⁻ and *Abd-B*⁻ is equivalent to that of a deletion of the entire BX-C such as *DfP9*. This suggests that these genes represent all of the functions of BX-C. Nevertheless, definitive proof requires a study of the phenotype of double and triple mutant combinations.

Subsidiary functions and compartments

The available evidence suggests that individual BX-C functions are used on a compartment-by-compartment basis. These functions are grouped genetically into three major complementation units defining three domains of higher hierarchical order; each

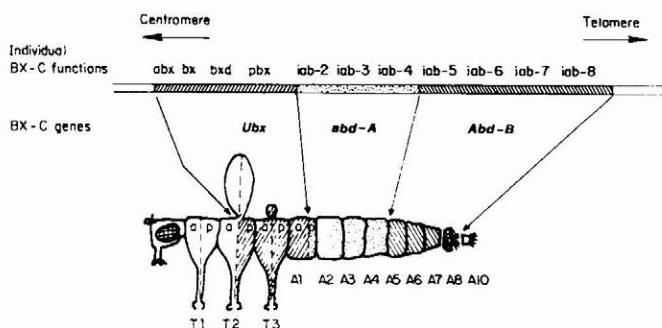


Fig. 4 Genetic and anatomical domains of the BX-C. The entire function of the BX-C extends from the anteroposterior boundary in T2 posteriorly to A8 (note that in T2p, the *Ubx* domain extends only to the ventral area, as *Ubx* mutations do not seem to affect the dorsal part¹⁹). Regions anterior or posterior to these boundaries such as head or analia (A10) do not require BX-C activity. However, as the location of this posterior boundary is based mainly on the phenotypes of escapers, it is conceivable that it could be more posterior. The BX-C is composed primarily of a number of individual functions *abx-bx-bxd*, ..., *iab-7*, *iab-8*. Some of these primary functions are well characterized, such as the thoracic series *abx-bx-bxd-pbx*¹⁻³, each determining the development of a specific anterior or posterior compartment. The abdominal series is less well characterized and it is not yet known whether these genes act within compartments or entire segments. For simplicity, we have assumed that there is an abdominal function per segment as proposed by Lewis^{2,3}, but it could be that there are two, one for the anterior and a second for the posterior compartment of each segment. We propose that all these individual functions can be grouped into three major genetic domains that correspond to our three complementation units: *Ubx*⁺, *abd-A*⁺ and *Abd-B*⁺. Each of these genes is responsible for the characteristic pattern of a specific anatomical domain. Except in the case of *Abd-B*, these domains do not correspond with the entire realm of action of the corresponding genes; they represent the area where the gene generates the morphological diversity. *Ubx*⁺ contains the *abx-bx-bxd-pbx* functions: the loss of the activity of *Ubx*⁺ results in the quadruple mutant combination. The *Ubx* mutants fail to complement any of the *abx*, *bx*, *pbx* and *bxd* mutations^{1,2,9}, but some combinations of these show complete or nearly complete complementation. In general, *abx* and *bx* mutants complement *bxd* and *pbx*, although strong *bx* mutants produce a partial inactivation of *pbx*⁺ function¹. Molecular analysis¹⁴ has shown that *abx*, *bx*, *bxd* and *pbx* pseudalleles are grouped in specific areas of the *Ubx* domain, suggesting a functional diversity within the *Ubx*⁺ gene. Note that, in contrast to other recent models^{3,6,12} we do not consider *Ubx*⁺ and *bxd*⁺ to be in the same hierarchical level; instead, we propose *bxd* and *pbx* to be part of the *Ubx* domain. *abd-A*⁺: we tentatively include here the *iab-2-4* functions because *abd-A*⁺ probably contains several individual functions of the type proposed by Lewis^{2,3}; the only evidence for such genes is the allele *iab-2^K* (ref. 4). The independent mutations *iab-3* and *iab-4* have not yet been detected. However, the fact that segment A3 is different from A2 and *iab-2* affects only A2 (ref. 4 and data reported here), suggests there must exist at least one function to generate the difference. Segments A3 and A4 are virtually identical, which raises the question as to whether there are different *iab-3*⁺ and *iab-4*⁺ functions. At present, it is not possible to answer this question; there is evidence, however, from the thoracic functions¹⁰ indicating that *bx*⁺ and *pbx*⁺ are duplicates of the same developmental function acting in different positions. A similar phenomenon could occur in the case of *iab-3*⁺ and *iab-4*⁺. All *abd-A* mutations affect several segments extending from A1p to at least A6 and often to A8. This conclusion is based on the embryonic and adult phenotypes of *abd-A^{M1}*, *abd-A^{M3}*, *T(2;3)P10* and *abd-A^{MX1}* and in the viable *trans* combinations of *abd-A^{M3}* with all the lethal alleles. *Abd-B*⁺: we include in this gene the *iab-5*⁺ to *iab-8*⁺ of Lewis^{2,3}. The boundary of function of *abd-A*⁺ and *Abd-B*⁺ is uncertain because segments A3 and A4 are morphologically indistinguishable. All our *Abd-B* mutants affect several segments; A5-A7 are always affected in hemizygous combinations and A8 is affected in most of them. Thus they all include several functions of the *iab* type. Some combinations of several *Abd-B* mutants with the allele *iab-8* described by Lewis³ sometimes affect the genitalia.

major domain comprises several compartments which have independent lineage and pattern. A single gene product, if expressed equally in all segments or compartments, cannot explain the diversity within the domain. Either each major gene contains several discrete functions, or it is expressed differentially in the various compartments to generate the specific patterns.

In the *Ubx* domain the phenotypes of the mutations *abx*, *bx*, *pbx* and *bxd* are parts of the *Ubx*⁻ phenotype affecting individual compartments; they further subdivide the *Ubx* domain and their function is responsible for compartment identity.

In the *abd-A* domain, the mutation *iab-2^K* also appears to define a similar restriction, itself part of the *abd-A* phenotype,

suggesting a similar situation to that of *Ubx*. In the *Abd-B* domain the phenotypic variation shown by the mutants may also be due to a functional diversity within the locus.

Anteroposterior compartment boundaries often appear to demarcate BX-C functions—the best-known example of this is the *Ubx* domain, limited at both ends by anteroposterior boundaries^{6,9-12}. The adjacent *abd-A* domain is limited anteriorly by the same boundary at A1. Within the *Ubx* domain, all of the individual functions (*abx*⁺, *bx*⁺, *pbx*⁺, *bxl*⁺) are demarcated by both anteroposterior compartment and segment boundaries. This is because each segment is made of two compartments and each function is specific for one compartment. At the time the BX-C genes become active, there is no obvious difference between a segment boundary or an anteroposterior

compartment boundary. At the blastoderm stage, the anterior and posterior compartment primordia are already segregated, as shown by clonal analysis^{22,23}. Indeed, the blastoderm embryo can be described as composed of a number of A-P-A-P-A-P-A-P-A primordia where every boundary separates an anterior from a posterior compartment. During subsequent development, segment boundaries become morphologically distinct, leading to the separation of segments as anatomical units, whereas A-P boundaries do not correspond to clearcut morphological landmarks. Initially, this difference does not exist and therefore BX-C genes do not discriminate a P-A (later segment) from a A-P (later compartment) boundary.

We thank Peter Lawrence and Pat Simpson for advice and the CAICYT for financial support.

Received 6 August; accepted 22 October 1984.

- Lewis, E. B. *Am. Zool.* **3**, 33-56 (1963).
- Lewis, E. B. *Nature* **276**, 565-570 (1978).
- Lewis, E. B. in *Developmental Biology Using Purified Genes* (eds Brown, D. D. & Fox, C. F.) 189-208 (Academic, New York, 1981).
- Kuhn, D. T., Woods, D. F. & Cook, J. L. *Molec. gen. Genet.* **181**, 82-86 (1981).
- Morata, G., Botas, J., Kerridge, S. & Struhl, G. *J. Embryol. exp. Morph.* **78**, 319-341 (1983).
- Struhl, G. *Nature* **308**, 454-457 (1984).
- Kerridge, S. & Morata, G. *J. Embryol. exp. Morph.* **68**, 211-234 (1982).
- Struhl, G. *Nature* **293**, 36-41 (1981).
- Morata, G. & Kerridge, S. *Nature* **290**, 778-781 (1981).
- Morata, G. & García-Bellido, A. *Wilhelm Roux's Arch. dev. Biol.* **179**, 125-143 (1976).
- Miñana, F. J. & García-Bellido, A. *Wilhelm Roux's Arch. dev. Biol.* **191**, 331-334 (1982).
- Hayes, P. H., Sato, T. & Denell, R. E. *Proc. natn. Acad. Sci. U.S.A.* **81**, 545-549 (1984).
- Lindsley, D. L. & Grell, E. H. *Genetic Variations of Drosophila melanogaster*, No. 627 (Carnegie Instn, Washington, 1968).

- Bender, W. *et al. Science* **221**, 23-29 (1983).
- Lewis, E. B. *Drosoph. Inf. Serv.* **55**, 207-208 (1980).
- Lewis, E. B. *Proc. 12th int. Conf. Genet. (Abstr.)* **12**, 96-97 (1968).
- Lawrence, P. A. & Johnston, P. *Cell* **36**, 775-782 (1984).
- Struhl, G. *J. Embryol. exp. Morph.* **76**, 297-331 (1983).
- Morata, G. & Kerridge, S. *Nature* **300**, 191-192 (1982).
- Lewis, E. B. *Am. Nat.* **89**, 73-89 (1955).
- Struhl, G. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7380-7384 (1982).
- Steiner, E. *Wilhelm Roux Arch. dev. Biol.* **180**, 9-30 (1976).
- Lawrence, P. A. & Morata, G. *Dev. Biol.* **56**, 40-51 (1977).
- Morata, G. & Ripoll, P. *Dev. Biol.* **42**, 211-221 (1975).
- Van der Meer, J. *Drosoph. Inf. Serv.* **52**, 160 (1977).
- Lohs-Schardin, M., Cremer, C. & Nüsslein-Volhard, C. *Dev. Biol.* **73**, 239-255 (1979).
- Kornberg, T. *Dev. Biol.* **86**, 363-372 (1981).
- Nöthiger, R., Dübendorfer, A. & Epper, F. *Wilhelm Roux's Arch. dev. Biol.* **181**, 367-373 (1977).
- Dübendorfer, K. & Nöthiger, R. *Wilhelm Roux's Arch. dev. Biol.* **191**, 42-55 (1982).
- Morata, G. *Am. Zool.* **22**, 57-64 (1982).

LETTERS TO NATURE

Two remarkable bright supernova remnants

P. A. Shaver*, C. J. Salter†‡, A. R. Patnaik‡,
J. H. van Gorkom§ & G. C. Hunt§

* European Southern Observatory, Karl-Schwarzschild-Strasse 2,
D-8046 Garching bei München, FRG

† National Radio Astronomy Observatory, 949 North Cherry Avenue,
Campus Building 65, Tucson, Arizona 85721-0655, USA

‡ Tata Institute of Fundamental Research Centre, Indian Institute of
Science Campus, PO Box 1234, Bangalore 560 012, India

§ National Radio Astronomy Observatory, PO Box 0, Socorro,
New Mexico 87801, USA

Statistical studies of supernova remnants (SNRs) indicate that the objects of high surface brightness are young¹. This is confirmed by measurements of angular expansion for a number of such remnants and by the secular decrease in their radio flux densities². Studies of these young SNRs are of particular interest as they can impose constraints on models of supernova outbursts and can reveal details of the initial interaction between the rapidly-expanding stellar material and the ambient interstellar medium³. Also, a high percentage of bright remnants belong to the comparatively-rare class of filled-centre SNRs. However, because of their small angular sizes, many have not been observed with sufficient resolution to determine detailed structures or, in some cases, even to verify an SNR classification. Between July 1981 and August 1984, we used the Very Large Array (VLA) of the National Radio Astronomy Observatory (NRAO) in its B, C and D configurations to clarify the morphology of several high-brightness sources from the Milne SNR catalogue⁴. Of these, the two sources G349.7+0.2 and G357.7-0.1 were found to be most unusual.

G349.7+0.2 is the third brightest galactic SNR, after Cas A and the Crab Nebula. No previous radio map was of sufficient resolution to reveal its structure, the SNR classification resting

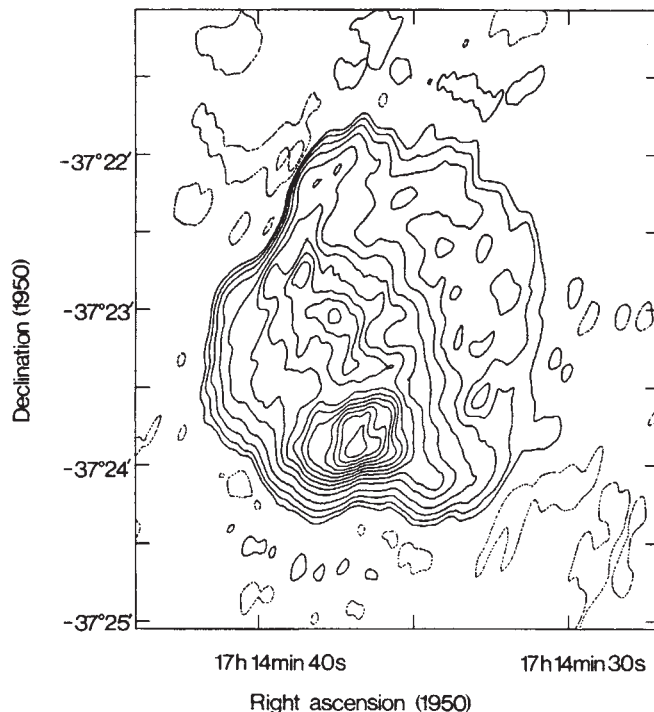


Fig. 1 1.4-GHz total intensity map of G349.7+0.2. The half-power beamwidth is 3.4×14.5 arc s at position angle -21° . Contour levels are $-5, -2, -1, 1, 2, 5, 10, 20, 30, \dots$ % of the peak brightness of 290 mJy per synthesized beam area.

on a non-thermal spectrum and an H I absorption distance of 18.3 ± 4.6 kpc⁵. No optical counterpart has been detected. The 1.4-GHz map from the VLA B array (Fig. 1) and the 4.9-GHz map from the C array are very similar, both containing about 85% of the total emission.



Fig. 2 Radiograph of the total intensity emission from G357.7-0.1 at 1.4 GHz. The half-power beamwidth is 3.8×10.9 arcs at position angle 19° .

The source is circular at the outermost contours (axial ratio ~ 0.9) and the peak emission lies near the edge (Fig. 1); in these respects it resembles the shell-type remnants rather than the more elliptical and centrally-peaked filled-centre SNRs³. Furthermore, its spectral index, α , is -0.47 ± 0.06 , close to the mean (-0.45) for shell-type remnants¹, but inconsistent with the flat spectra ($-0.3 < \alpha < 0$) found for filled-centre SNRs⁶. However, it is unusual for a shell-type remnant insofar as it lacks the

usually-prominent central cavity and ring-like appearance. There are, nevertheless, indications of partial shell structure, particularly in the lower contours to the north-west. The strong north-south emission ridge, located somewhat east of centre and merging with this shell structure at both ends, probably represents a non-uniform distribution of emission within the shell, rather than a high central volume emissivity as is found in filled-centre remnants such as the Crab Nebula. As such, this would be evidence for strong interaction with the interstellar medium at an early stage in the life of this young SNR.

G357.7-0.1 is the sixth brightest object in the Milne catalogue⁴. It has a relatively-steep radio spectral index, -0.53 ± 0.02 ; the only reliable distance information is a limit of > 6 kpc from H I absorption measurements⁷ and, again, no optical counterpart has been found. This is the closest (catalogued) SNR to the galactic centre direction and it could be located in the nuclear disk itself. Lower-resolution radio observations show extended emission aligned at position angle 65° , with the peak brightness offset to the south-west^{8,9}. Even its classification as an SNR has been questioned, some authors noting its similarity at this resolution to extragalactic head-tail sources^{6,10}; others have felt that, if it is an SNR, it is probably of the filled-centre variety despite the steep spectral index^{11,12}.

Figure 2 is a grey-scale representation of G357.7-0.1 from the combined VLA B- and C-configuration data at 1.4 GHz, and a contour map of the same data is shown in Fig. 3. These maps contain a flux density of about 35 Jy, somewhat greater than expected from existing pencil beam measurements (29 Jy). Figures 2 and 3 reveal this object to be one of the most remarkable radio sources yet mapped, with its unique structure of filamentary arcs. Although no other known SNR, either shell-type or filled-centre, has a similar appearance, the morphology would be even more abnormal for an extragalactic source. It is almost certainly galactic.

Measurable linear polarization, typically 10%, is found over most of G357.7-0.1 at 4.9 GHz and, in Fig. 4, vectors representing this polarized intensity are superimposed on the total-intensity contours from the D-array observations at 4.9 GHz. The vectors appear well ordered across the entire source, implying a well-ordered intrinsic magnetic field. There is no detailed correlation between the polarization and the filamentary structure. At present, we have insufficient information on the Faraday

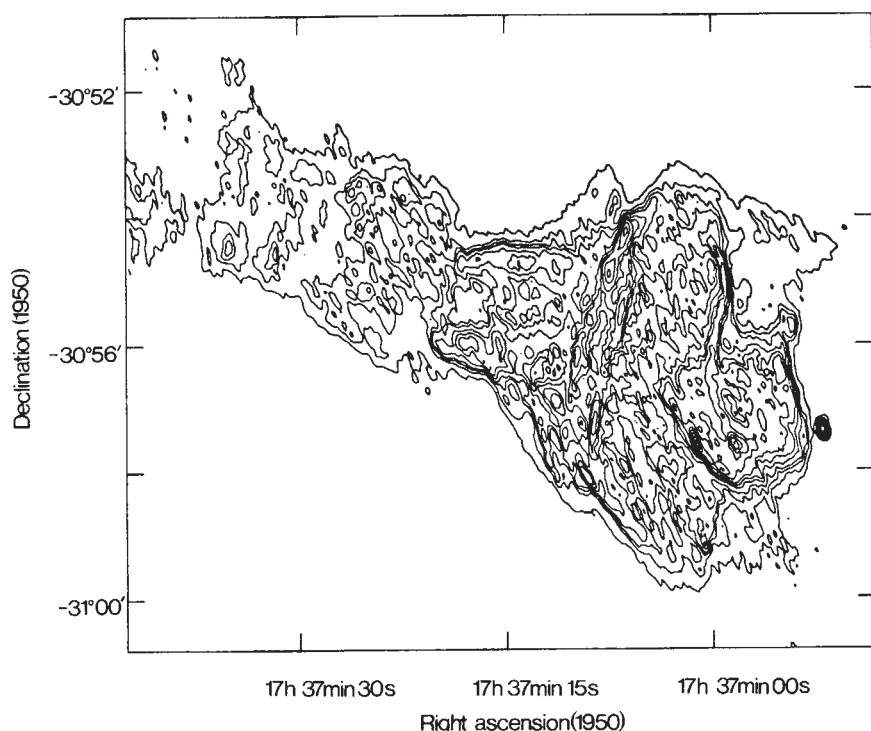


Fig. 3 1.4-GHz total intensity map of G357.7-0.1. The half-power beamwidth is 3.8×10.9 arcs at position angle 19° . Contour levels are 3, 6, 9, 12, 18, 24, 30, 42, 54, 66, 78, 90, 102, 114 mJy per synthesized beam area. No correction has been applied for the primary beam response of the individual antennas (half-power beamwidth ~ 30 arc min).

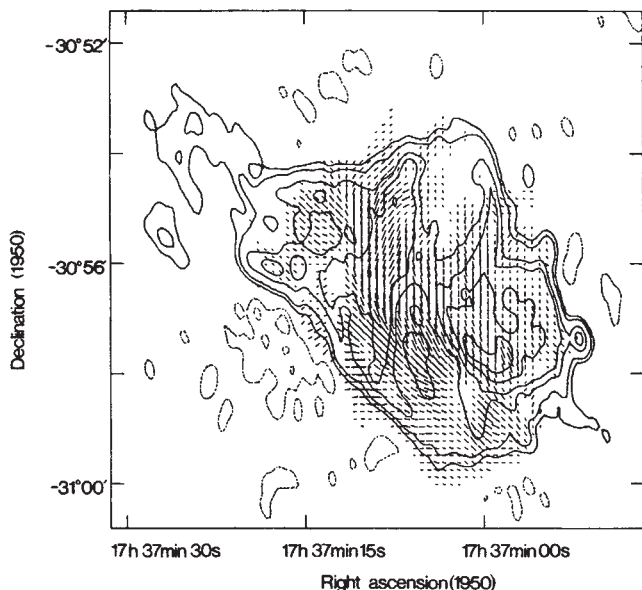


Fig. 4 The distribution of linear polarization on G357.7-0.1 at 4.9 GHz. The half-power beamwidth is 12×26 arc s at position angle 9° . Contour levels are -10, -5, 5, 10, 30, 50, 100, 150 mJy per synthesized beam area. A vector of length 10 arc s represents 4.7 mJy per synthesized beam area of polarized intensity. The orientation of the vectors gives the position angle of the E vector of the polarized emission. Polarized intensities of <0.9 mJy per synthesized beam area are not plotted. No correction has been applied for the primary beam response of the individual antennas (half-power beamwidth ~ 9 arc min) and consequently the flux density in the contours is reduced to $\sim 70\%$ of that expected from pencil beam measurements.

rotation of the polarized emission to draw any conclusions on magnetic field directions within the object.

The filaments are very thin, with deconvolved widths of ~ 5 arc s, implying length/width ratios of up to 60. Their brightness, relative to the more smoothly-distributed emission, is far higher than in the Crab Nebula; indeed they are reminiscent of the arc features recently resolved near the galactic centre¹³. In these respects, this object is unique as an SNR.

The three-dimensional structure is difficult to elucidate. There is some suggestion of a structural axis, corresponding roughly to the major axis of the source at position 65° and around which one may discern partial helical or cylindrical structures, the prominent filaments possibly delineating a 'waist'. This could conceivably be construed as evidence for a predominantly-equatorial supernova outburst, as has been suggested¹⁴ may result from the collapse of a rotating massive star. Any helical structure is also reminiscent of the precessing object SS433, although the scale is completely different, and G357.7-0.1 is quite unlike the SNR W50 surrounding SS433^{15,16}. Alternatively, the 'head-tail' character of the structure illustrated in Fig. 3 may indicate some sort of evolution from east to west, although there is no evidence that the spectral index changes markedly across the source. We note in this regard the intriguing presence of a point source near the major axis of G357.7-0.1, just beyond the brighter end. It has a flux density of 30 ± 10 mJy at 1.4 GHz and 48 ± 14 mJy at 4.9 GHz, indicating a flat or inverted spectrum. Although these properties are typical of compact extragalactic sources (and there is no obvious link between this source and the extended object), further investigation would be of interest.

The sharpness of the filaments and the presence of the point source as a phase reference will make it possible to map proper motions in G357.7-0.1. A lateral velocity of $1,000 \text{ km s}^{-1}$ at a distance of 10 kpc would produce a proper motion of 0.1 arc s in 5 yr. Such proper motion studies will help to elucidate the geometry, nature and origin of this object.

At present, we can only say that the filamentary morphology, non-thermal emission and polarization of G357.7-0.1 seem to indicate that it is an SNR. It does not fall clearly into filled-centre or shell type, but the ordered nature of the main filaments, the off-centre position of the radio peak and the steep radio spectrum all suggest a closer affinity with the shell-type variety. However, the unusual properties of G357.7-0.1 suggest that its origin is as unique as its appearance.

The NRAO is operated by Associated Universities, Inc. under contract with the NSF.

Received 26 October; accepted 28 November 1984.

1. Clark, D. H. & Caswell, J. L. *Mon. Not. R. astr. Soc.* **174**, 267-305 (1976).
2. Strom, R. G., Goss, W. M. & Shaver, P. A. *Mon. Not. R. astr. Soc.* **200**, 473-487 (1982).
3. Shaver, P. A. *Astr. Astrophys.* **105**, 306-312 (1982).
4. Milne, D. K. *Aust. J. Phys.* **32**, 83-92 (1979).
5. Caswell, J. L., Murray, J. D., Roger, R. S., Cole, D. J. & Cooke, D. J. *Astr. Astrophys.* **45**, 239-258 (1975).
6. Weiler, K. W. & Shaver, P. A. *Astr. Astrophys.* **70**, 389-398 (1978).
7. Radhakrishnan, V., Goss, W. M., Murray, J. D. & Brooks, J. W. *Astrophys. J. Suppl.* **24**, 49-121 (1972).
8. Slee, O. B. & Dulk, G. A. *IAU Symp.* **60**, 347-351 (1974).
9. Caswell, J. L., Clark, D. H. & Crawford, D. F. *Aust. J. Phys. Suppl.* **37**, 39-56 (1975).
10. Weiler, K. W. & Panagia, N. *Astr. Astrophys.* **90**, 269-282 (1980).
11. Lockhart, I. A., Goss, W. M., Caswell, J. L. & McAdam, W. B. *Mon. Not. R. astr. Soc.* **179**, 147-152 (1977).
12. Caswell, J. L., Haynes, R. F., Milne, D. K. & Wellington, K. J. *Mon. Not. R. astr. Soc.* **190**, 881-889 (1980).
13. Yusef-Zadeh, F., Morris, M. & Chance, D. *Nature* **310**, 557-561 (1984).
14. Bodenheimer, P. & Woosley, S. E. *Astrophys. J.* **269**, 281-291 (1983).
15. Hjellming, R. M. & Johnston, K. J. *Nature* **290**, 100-107 (1981).
16. Geldzahler, B., Pauls, T. & Salter, C. J. *Astr. Astrophys.* **84**, 237-244 (1980).

A new class of nonthermal radio sources

R. H. Becker* & D. J. Helfand†

*University of California at Davis, Department of Physics, Davis, California 95616, USA

†Columbia Astrophysics Laboratory, 538 W. 120th Street, New York, New York 10027, USA

Current radio catalogues^{1,2} list some 145 sources as galactic supernova remnants which are included primarily on the basis of their exhibiting a nonthermal radio spectrum, evidence for extended emission and being located within a few degrees of the galactic plane. Detailed maps including spatially-resolved spectral information and polarimetric data have been obtained for about half of the catalogued objects, primarily those lying within ~ 5 kpc of the Sun, and have revealed three classes of remnants: shell-like sources, Crab-like synchrotron nebulae and composite remnants^{3,4} containing a central Crab-like component surrounded by a shell. While mapping many of the remnants not well studied to date, we have discovered at least two nonthermal extended objects which defy classification in any of these categories. The two sources exhibit a marked axial symmetry with evidence for a compact component at one extreme of the axis. Their diffuse emission includes bright filamentary structures superimposed on a lower surface brightness component which gradually fades, becoming undetectable at the end of the axis opposite the compact source. The mean spectral indices of these sources are intermediate between the shell- and Crab-like remnants. We propose here that these objects are representative of a new class of nonthermal radio sources; elsewhere in this issue⁵, we suggest that such sources are formed through the emission of relativistic particles produced by high-velocity accreting binary systems.

The radio sources G357.7-0.1 and G5.3-1.0 were observed with the National Radio Astronomy Observatory Very Large Array (VLA) in the spring and summer of 1984 at 20 and 6 cm utilizing the B-C and C-D hybrid configurations, preferable for sources at southern declinations as they produce a more circular synthesized beam. Although, by using two different arrays for the two wavelengths, we have achieved comparable resolution

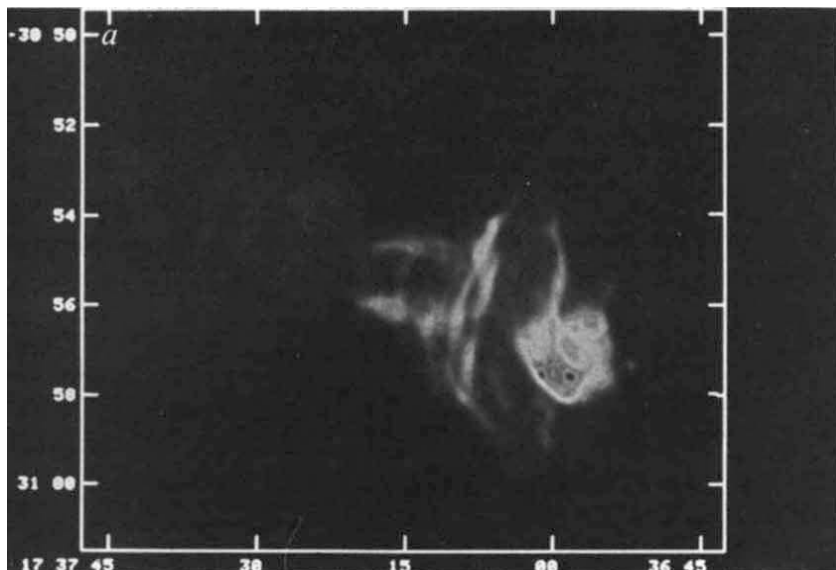
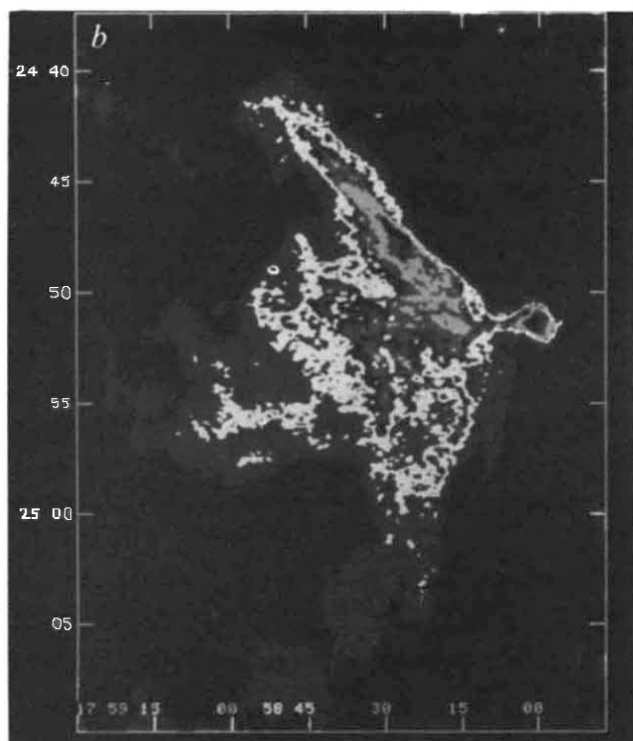


Fig. 1 *a*, A radiograph of the non-thermal galactic radio source G357.7-0.1 at 20 cm obtained with the VLA. *b*, A radiograph of the non-thermal galactic radio source G5.3-1.0 at 20 cm obtained with the VLA.



in Fig. 2a; there was no indication of polarized emission at 20 cm.

The 6-cm image, shown as a contour map in Fig. 2b, reveals the same structure in the western end of the source, but fails to detect the emission from the eastern half, partly because of its lower sensitivity in detecting extended features. In addition, the 6-cm emission was linearly polarized in places (Fig. 3), particularly from the second and largest ring structure.

Comparing Fig. 2a and b provides information about the spectrum as a function of position of the source (but note that this procedure is somewhat uncertain, given that the 6-cm data are less sensitive to extended emission). Starting at the west, we have fitted single gaussian models to the compact source of emission and, taking the average of the two wavelengths, we locate the source at RA (1950) = 17 h 36 min 52.11 s and dec. (1950) = $-30^{\circ}57' 21.65''$. The fits at both frequencies result in a finite size estimate of 3-4 arc s, but the source sits on top of a nonuniform background resulting in a formal uncertainty in the size of ± 2 arc s, and thus it could be point-like. The flux is estimated at 60 ± 5 mJy and 44 ± 5 mJy at 20 and 6 cm respectively, implying a spectral index of ~ 0.25 ; in comparison, the spectral index for the hot spots in the southern end of the first ring is ~ 0.5 , similar to the spectrum for the integrated total intensity of ~ 0.45 (ref. 5).

A contour plot of the 20-cm radiograph of the second source, G5.3-1.0 (Fig. 4), is superficially dissimilar to the G357.7-0.1 image, but shares many of its characteristics, including axial symmetry, a gradient in surface brightness from west to east and a compact source at one extreme of the axis; the only significant difference is a somewhat flatter integrated spectral index of 0.2 (ref. 5). The image of G5.3-1.0 shows a source 20 arc min in length with a leading compact object on the western boundary behind which extend several bubbles and arcs of emission followed by a long low-surface-brightness tail. This image is the result of 20-cm data combined from both the B-C and C-D arrays; the shorter spacings were necessary to obtain a more complete picture of the extended structure. Comparable short spacings were unavailable for the 6-cm observations (not shown) and only the highest surface brightness features are visible in that image, but, as usual, the structure of the polarization is characterized by smaller scales and is mapped quite well at 6 cm. The polarized intensity and E-vector orientations are depicted in Fig. 5, which shows that most of the polarized emission arises in the northern 'wing' of the source, is remarkably uniform and everywhere perpendicular to the source boundary. Assuming little Faraday rotation, an assumption open to question, the magnetic field lines are thus parallel to the source boundary.

In summary, G5.3-1.0 and G357.7-0.1 are similar in their

in the resulting images, the correspondence is not exact, as mutual shadowing of the most closely spaced antennas occurs more frequently in the compact C-D configuration. As a result, the 6-cm image contains somewhat less information about extended structure; the smaller primary beam at 6 cm also reduces the sensitivity to features on the largest angular scale (~ 30 arc min). A total of ~ 40 min of data were accumulated for each source at each of the two frequencies. Maps were then produced from the raw data and subsequently cleaned using standard algorithms.

The radiograph of the 20-cm brightness distribution, shown in Fig. 1, best exemplifies the unique nature of G357.7-0.1. The immediate impression is that of a number of smoke rings of varying dimension, concentrically placed along a common axis; we can trace emission along this axis for ~ 12 arc min, the intensity decreasing monotonically from west to east suggesting a temporal sequence. We also note the compact source of emission located west of the westernmost ring and lying along the same axis. The same data are represented in a contour map

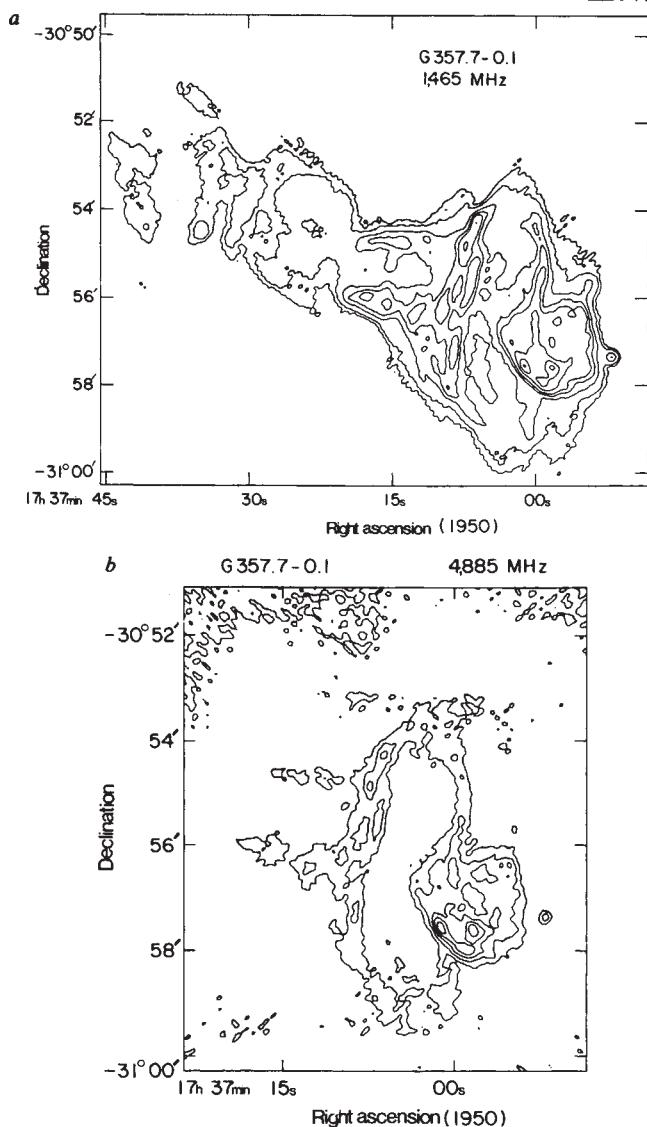


Fig. 2 A contour plot of the radio image of G357.7-0.1 at 20 cm (a) and 6 cm (b). The contour levels correspond to a: 2, 5, 15, 25, 35, 55, 75 and 95% of a peak flux of $0.66 \text{ Jy beam}^{-1}$ and b: 3, 5, 7, 9 and 11% of a peak flux of $0.76 \text{ Jy beam}^{-1}$.

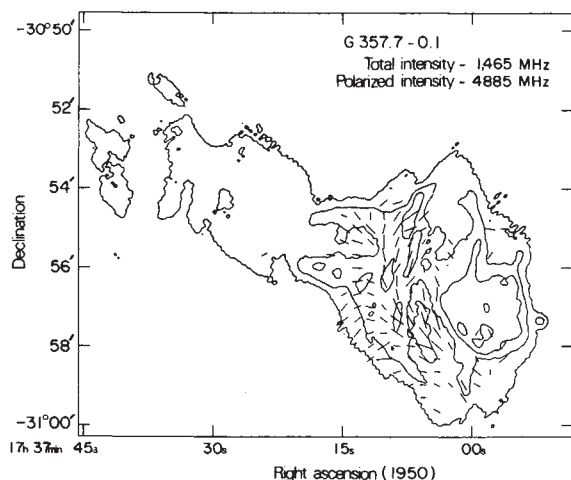


Fig. 3 The linear polarization vectors for G357.7-0.1 from the 6-cm polarization map overlaid on the 20-cm contour plot. The orientation of the straight lines represents the direction of the E-vector of the received radiation whereas the length of each line is proportional to the polarized intensity at that location (1 arc s = $0.09 \text{ mJy per beam}$).

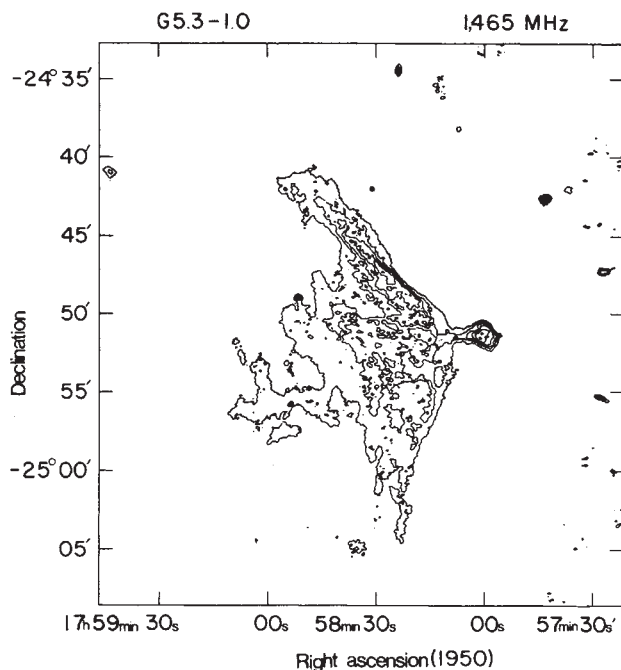


Fig. 4 A contour plot of the radiograph of G5.3-1.0 as observed at 20 cm. Contour levels correspond to 3, 5, 7, 9, 12 and 15% of a peak flux of 78 mJy beam^{-1} (the peak flux in the map is outside the region shown).

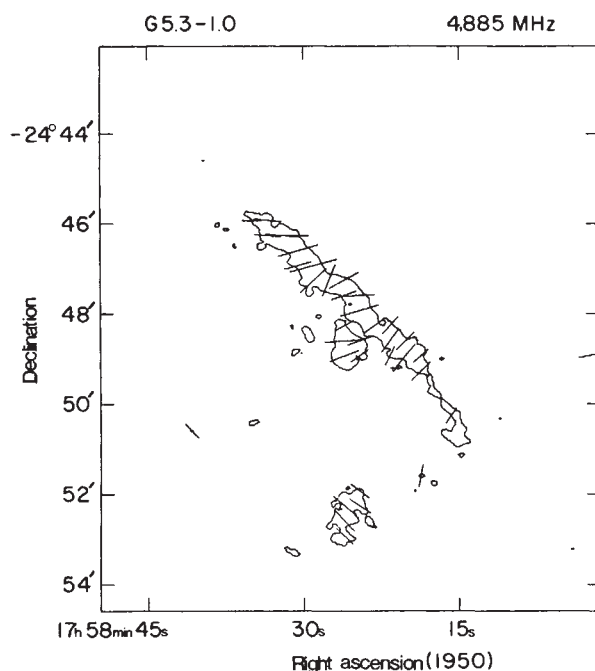


Fig. 5 Polarized intensity along the northern wing of G5.3-1.0 as observed at 6 cm; polarization vectors are as in Fig. 3. The contour level is $0.7 \text{ mJy beam}^{-1}$; a line 10 arc s long corresponds to 0.3 mJy per beam area of polarized intensity.

detailed morphology, yet differ markedly from any previously observed radio sources. Both objects are clearly non-thermal with significant linear polarization, both have axial symmetry with a progression of features from compact at one extreme to diffuse at the other, and both images suggest episodic outbursts of energy. The failure to recognize these objects previously is, in part, because of the complacency with which all non-thermal radio sources near the galactic plane have been accepted as supernova remnants, but also owing to the lack of high resolution images for southern sources. Although we do not expect to find

an abundance of these objects, we would not be surprised were additional sources of this type to be discovered among those presently catalogued as supernova remnants.

The National Radio Astronomy Observatory is operated by Associated Universities, Inc., under contract with the NSF. R.H.B. thanks NSF for support under grant AST-8313685 and D.J.H. thanks the US Air Force Office of Scientific Research (AFOSR 82-0014) and the Alfred P. Sloan Foundation under its Research Fellow program. This is contribution no. 277 of the Columbia Astrophysics Laboratory.

Received 1 November; accepted 28 November 1984.

1. Milne, D. E. *Aust. J. Phys.* **32**, 83-92 (1979).
2. Green, D. A. *Mon. Not. R. astr. Soc.* **209**, 449-478 (1984).
3. Becker, R. H. & Helfand, D. J. *Astrophys. J.* **283**, 154-157 (1984).
4. Weiler, K. W. *IAU Symp.* No. 101, 299-320 (1984).
5. Helfand, D. J. & Becker, R. H. *Nature* **313**, 118-119 (1984).
6. Green, A. J. *Astr. Astrophys. Suppl.* **18**, 267-307 (1974).

Origin of the new axisymmetric radio sources

D. J. Helfand* & R. H. Becker†

* Columbia Astrophysics Laboratory, 538 W. 120th Street, New York, New York 10027, USA

† University of California at Davis, Department of Physics, Davis, California 95616, USA

In the accompanying letter¹, we have reported the discovery of two nonthermal radio sources lying near the galactic plane, which have surface brightness distributions radically different from those of the supernova remnants (SNRs) with which they had been formerly classified. Their marked axial symmetry, associated compact components, steep surface-brightness gradients and polarization structure are unlike those seen in any previously-identified class of radio emitters. Here, we discuss some of their physical properties (distances, ages, energetics) and describe briefly a scenario in which the observed radiation is produced by electrons generated in a relativistic outflow from high-velocity accreting binary systems.

The radio emission from shell-type SNRs is understood to arise from particles accelerated at the shock front formed as a supernova blast-wave expands into the ambient medium. On achieving energies of 10^9 – 10^{11} eV, the electrons produce synchrotron radiation in the region of compressed magnetic fields behind the shock ($B \sim 10^{-5}$ G). The lifetime of a source ($\leq 10^5$ yr) is limited to the age at which the SN-shock velocity falls to a value (~ 20 km s⁻¹) comparable with the characteristic turbulent velocity of the interstellar medium, whereas the evolution of the radio luminosity of the remnant is determined primarily by adiabatic-expansion losses undergone by the synchrotron electrons.

In the Crab-like remnants, the radio-emitting electrons are produced by a central pulsar, deriving their energy, directly or indirectly, from the rotational-kinetic-energy loss of this source. They radiate in a magnetic field ($\sim 5 \times 10^{-4}$ G in the case of the Crab) which is also generated by the pulsar and confined (at least in some sources) by the shell of supernova ejecta. As the synchrotron lifetimes of the particles are much greater than the characteristic energy-loss time scale for the pulsar ($\sim 10^3$ yr), and as the acceleration mechanism extends to much higher energies ($\gamma \geq 10^7$), providing a source of radio particles evolving downward in energy from that of X-ray and optical synchrotron electrons, the radio-luminosity evolution of these remnants is non-monotonic. A successful self-consistent model has yet to be produced, but a number of indirect arguments^{2,3} suggest that the lifetimes of the Crab-like sources are considerably shorter ($\sim 10^4$ yr) than those of shell-like objects. The newly-emerging class of composite remnants^{4,5}, in which both radio shells and Crab-like cores are visible, probably represent a relatively brief

period in the history of pulsar-driven remnants in which the emission from both components is comparable (that is, within a factor $< 10^2$).

The mean radio spectral index of the shell-type remnants is ~ 0.5 , whereas for the Crab-like nebulae, α ranges from 0.0 to 0.2. Both classes show roughly circular symmetry with little spatial variation in spectral index across the face of the source. For all remnants exhibiting a compact object⁶ (for example, a radio pulsar or a point-like X-ray source), it is found inside the boundary of the emission, usually quite close to the centre. This is to be expected as the typical velocities of normal stars, pulsars and accreting binaries are all $\leq 10^2$ km s⁻¹, whereas the initial expansion velocity of the SNR is $\sim 10^4$ km s⁻¹ and the mean shock speed remains $> 10^2$ km s⁻¹ for $> 10^4$ yr; thus in the lifetime of either remnant class, a compact object will not outrun the shock front.

Clearly, the two extended nonthermal radio sources¹ do not fit the description of either type of SNR, their symmetry being axial rather than circular and their spectral indices falling between those of the two remnant classes; the highest surface-brightness features and possible point-source components lie at one extreme of the axis, outside the main emission region. Thus, we believe that these objects represent a new class of diffuse non-thermal radio source.

Their large angular size and proximity to the galactic plane argues strongly that they are indeed galactic objects (at a redshift of $z \sim 0.1$, the linear size of G5.3-1.0 would be ~ 3 Mpc). Little information is available on the distance to either source, although a hydrogen-absorption-line map of G357.7-0.1 (ref. 7) reveals evidence for absorption arising in the expanding 3-kpc arm of the Galaxy, placing the source at a distance ≥ 7 kpc from the Sun. Unfortunately, the low galactic longitude of both objects precludes the possibility of obtaining more accurate distances by this method, but suggests on statistical grounds a distance of ~ 10 kpc which we shall use, expressing our results in terms of the scaling factor d_{10} .

The angular size of the two sources is similar; from end to end, the axes span 12 and 20 arc min, implying linear extents of 36 and 60 d_{10} pc for G357.7-1.0 and G5.3-0.1 respectively. Their total radio luminosities are obtained from single-dish continuum measurements¹; using power law spectral indices determined from the ensemble of these measurements and integrating from $\nu_l = 10^7$ to $\nu_u = 10^{11}$ Hz, we obtain values of $L_R(\text{G357}) \sim 1 \times 10^{35} d_{10}^2$ and $L_R(\text{G5.3}) \sim 2 \times 10^{35} d_{10}^2$ erg s⁻¹. Thus, the total energy in radiating electrons can be approximated as

$$E_e \approx \frac{10^{12} L_R}{B^{3/2}} \left(\frac{\nu_u^{0.5-\alpha} - \nu_l^{0.5-\alpha}}{\nu_u^{1-\alpha} - \nu_l^{1-\alpha}} \right) \frac{2\alpha+2}{2\alpha+1} d_{10}^2 \quad (1)$$

where B is the mean magnetic field⁸. Adopting the interstellar value of $B \sim 10^{-5.5}$ G, we find $E_e \sim 3\text{--}20 \times 10^{49} d_{10}^2$ erg. Therefore, to produce radio emission over the observed band, the individual particles must have $10^3 \leq \gamma \leq 10^5$.

Without a detailed model for the particle acceleration mechanism, it is difficult to estimate how much additional energy goes into accelerating the unseen protons. In extragalactic radio sources, internal Faraday depolarization indicates a proton density in the jet flows implying $E_p \sim 10^2 E_e$, whereas in the Crab Nebula, limits on the dynamical pressure exerted by the relativistic particle population require $E_p \leq E_e$. Although a higher magnetic field strength could clearly lower the inferred total particle energy, it could only do so if the energizing source could provide sufficient additional energy to produce this enhanced field. For a volume of $\sim 20 d_{10}$ pc in radius (the mean value for the emitting regions), the interstellar field energy is $\sim 10^{49}$ erg, comparable with the particle energy calculated above. Furthermore, although the field energy grows as B^2 , the required particle energy declines only as $B^{3/2}$, resulting in a total energy in the field and particles $E_T \approx 10^{50} (B/10^{-5.5})^{1/2}$. Thus, we adopt 10^{50} erg as a lower limit to the energy input required to produce the observed emission.

Although the origin of this energy is uncertain, the overall morphology of these sources seems to suggest the passage of

an energizing agent, along an axis bisecting the diffuse emission, whose current activity is displayed in the compact components at the western end of each source. A reasonable transverse velocity of $V \sim 100 \text{ km s}^{-1}$ leads to a source age of $\tau \sim 5 \times 10^5$ (d_{10}/V_{100}) yr, which is comfortably shorter than the particles' synchrotron lifetimes if $B < 10^{-4} \text{ G}$. (In fact, if the fading of the emission at the eastern ends of the sources is caused by synchrotron losses, adiabatic losses are negligible, the field is just $\sim 10^{-4} \text{ G}$ and the total energy input is $\sim 10^{50.5} \text{ erg}$.) If the energy input has been approximately constant over this interval, the luminosity of the energizing source in high-energy particles must be $\geq 10^{37} \text{ erg s}^{-1}$.

Total luminosities exceeding 10^{50} erg over $> 10^{5.5} \text{ yr}$ occur in only two classes of galactic objects: massive ($> 5 M_{\odot}$) main sequence stars and accreting binary systems containing a neutron star or a black hole. Isolated neutron stars can indeed supply the total energy required (and have the added virtue of being known producers of relativistic electrons), but the distribution of pulsar periods and period derivatives strongly indicates that they maintain spin-down luminosities of $> 10^{37} \text{ erg s}^{-1}$ for $\leq 10^4 \text{ yr}$. There is no evidence that OB stars produce relativistic winds and we can think of no plausible mechanism by which this may occur. Nevertheless, a scenario involving a travelling binary system is both energetically and astrophysically feasible.

Several accreting binaries containing compact objects have been implicated in the production of high-energy particles. Sco X-1 is known to have a pair of radio lobes separated from the central source by $\sim 0.1 \text{ pc}$ (ref. 9). SS433 produces powerful jets of relativistic material¹⁰ (containing 10^{38} – $10^{42} \text{ erg s}^{-1}$), whereas the X-ray binaries Her X-1 and Vela X-1 have recently been reported as very-high-energy and/or ultra-high-energy γ -ray emitters^{11,12}, implying copious production of ultra-relativistic particles (~ 1 – 10% of the accretion energy). Cyg X-3 is also a confirmed 10^{11} – 10^{12} eV ($\gamma \sim 10^7$) γ -ray source^{12,13} which, over the past 10 yr, has experienced several large radio outbursts injecting $\geq 10^{45}$ – 10^{50} erg into an expanding shell of radio emitting plasma possibly via jets with an expansion speed of $\sim 0.35 c$ (ref. 14). Periodic flares with an energy input of $\leq 10^{43} \text{ erg}$ every $\sim 5 \text{ h}$ ($\sim 10^{51} \text{ erg per } 10^5 \text{ yr}$) have been reported recently¹⁵. Finally, on general theoretical grounds, it has been argued¹⁶ that compact objects surrounded by accretion disks should be expected to produce relativistic particle beams emerging along the disk rotation axis.

These binary systems are all strong X-ray sources with luminosities in the 0.2–20 keV band of 10^{36} – $10^{38} \text{ erg s}^{-1}$; significant unobservable extreme ultraviolet emission is also expected, leading to total accretion luminosities substantially exceeding the $\sim 10^{37} \text{ erg s}^{-1}$ that we require in high-energy particle flux. We predict, then, that compact X-ray sources should be found coincident with the leading (western) end of each of our new sources, having luminosities $\geq 10^{37} \text{ erg s}^{-1}$ but heavily absorbed at energies below 2–3 keV owing to the large column density of material along the line of sight to the galactic centre ($N_{\text{H}} > 10^{23} \text{ cm}^{-2}$).

Although the overall morphology and energetics of these two sources are consistent with the above scenario, the detailed structure of the emission also deserves comment. The simplest interpretation is that the varying distance from the axis at which emission is seen, and the filamentary structure of both the total intensity and the polarized flux, result from variations in the energy injection rate and/or the ambient magnetic field into which the particles are injected (recalling that mean interstellar field values of $3 \times 10^{-6} \text{ G}$ are sufficient to explain the emission). In particular, the 20-pc-long filament near the leading edge of G5.3–1.0 with its remarkably straight magnetic field is reminiscent of those recently reported¹⁷ near the galactic centre. Although the origin of the latter filaments remains unclear, in the case of G5.3–1.0, it is possible that the outflowing beam 'combs out' the ambient field and/or that the pre-existing field is amplified by shear stresses at the boundary of the flow, as is thought to occur in extragalactic jet sources. Higher sensitivity polarimetric mapping, together with a more detailed develop-

ment of source models, will be necessary before the origin of the fine structure in the morphology of these sources becomes clear.

R.H.B. thanks the NSF for support under grant AST-8313685 and D.J.H. thanks the US Air Force Office of Scientific Research (AFOSR 82-0014) and the Alfred P. Sloan Foundation under its Research Fellow program. This is contribution no. 281 of the Columbia Astrophysics Laboratory.

Received 1 November; accepted 28 November 1984.

1. Becker, R. H. & Helfand, D. J. *Nature* **313**, 115–118 (1984).
2. Weiler, K. W. & Pangia, N. *Astr. Astrophys.* **90**, 269–282 (1980).
3. Reynolds, S. & Chevalier, R. Preprint (National Radio Astronomy Observatory, 1984).
4. Becker, R. H. & Helfand, D. J. *Astrophys. J.* **283**, 154–157 (1984).
5. Helfand, D. J. & Becker, R. H. *Proc. Conf. Crab-like SNR USA* (in the press).
6. Helfand, D. J. & Becker, R. H. *Nature* **307**, 215–221 (1984).
7. Radhakrishnan, V., Goss, W. M., Murray, J. D. & Brooks, J. W. *Astrophys. J. Suppl.* **24**, 49–121 (1972).
8. Lang, K. R. *Astrophysical Formulae* 39 (Springer, New York, 1974).
9. Fomalont, E. B., Geldzahler, B. J., Hjellming, R. M. & Wade, C. M. *Astrophys. J.* **275**, 802–807 (1983).
10. Margon, B., Grandi, S. & Downes, R. A. *Astrophys. J.* **241**, 306 (1980).
11. Dowthwaite, J. C. et al. *Nature* **309**, 691–693 (1984).
12. Protheroe, R. J., Clay, R. W. & Gerhardt, P. R. *Astrophys. J. Lett.* **280**, L47–L50 (1984).
13. Lamb, R. C., Godfrey, C. P., Wheaton, W. A. & Turner, T. *Nature* **296**, 543–544 (1982).
14. Geldzahler, B. J. et al. *Astrophys. J. Lett.* **273**, L65–L69 (1983).
15. Molnar, L. A., Reid, M. J. & Grindlay, J. E. *Nature* **310**, 662–665 (1984).
16. Rees, M. J. *A. Rev. Astr. Astrophys.* **22**, 471–506 (1984).
17. Yusef-Zadeh, F., Morris, M. & Chance, D. *Nature* **310**, 557–561 (1984).

An extended X-ray low state from Hercules X-1

A. N. Parmar*, W. Pietsch†, S. McKechnie‡, N. E. White*, J. Trümper†, W. Voges† & P. Barr*

* EXOSAT Observatory, European Space Operations Centre, Robert Bosch Strasse 5, 6100 Darmstadt, FRG

† Max-Planck-Institut für Extraterrestrische Physik, D-8046 Garching bei München, FRG

‡ Laboratory for Space Research Leiden, Wassenaarseweg 78, PO Box 9504, 2300 RA Leiden, The Netherlands

Hercules X-1 exhibits a 35-day cycle in its X-ray intensity^{1,2} in addition to its pulsar rotational and orbital periodicities of 1.24 s and 1.7 days respectively. The effects of X-ray heating on the companion are visible throughout the 35-day cycle³, suggesting that the observed intensity modulation is caused by periodic blocking of the line of sight to the pulsar. We report here observations made with the EXOSAT Observatory between 1983 June and August that failed to detect the expected 35-day variation in X-ray intensity, although low-level extended X-ray emission was seen. Models where the obscuring of the pulsar is caused by a precessing tilted accretion disk have been proposed^{4–9}, although the physics of the tilted disk are not clear¹⁰. Alternatively, wave-like thickenings of the accretion disk have been proposed¹¹ to account for the observed modulation. Our EXOSAT observations suggest that a temporary change in the disk structure may have occurred such that the disk was in the line of sight throughout.

Hercules X-1 (Her X-1) was observed by the European Space Agency's EXOSAT Observatory (see ref. 12 for additional details) on 1983 June 28, ~ 37 days after the Tenma X-ray satellite observed that it was close to its maximum high-state intensity¹³. During the 35-day cycle, Her X-1 normally has a high-state lasting for ~ 10 days with a 2–6 keV intensity of ~ 0.1 , seen from the Crab Nebula (100 mCb), and a ~ 5 -day secondary high-state with peak intensity of $\sim 30 \text{ mCb}$ that occurs $\sim 180^\circ$ out-of-phase with the primary high-state. For the remainder of the 35-day cycle, referred to as the low state, Her X-1 is much fainter at $\sim 5 \text{ mCb}$ ^{1,2,14–16}. In the EXOSAT medium energy (ME) proportional counter array¹⁷, Her X-1 was detected only with a 2–6-keV count rate of $\sim 5 \text{ mCb}$, similar to that obtained in previous low-state measurements. Subsequent observations, approximately evenly spaced every 10 days over the following 70 days failed to detect a high-state (Table 1). A recent EXOSAT

Table 1 EXOSAT observations of Hercules X-1

	Time (UT) 1983		1.7-day* phase	35-day† cycle phase	ME 2-10 keV‡ flux (10^{-11} erg cm $^{-2}$ s $^{-1}$)
	Start h	Duration h			
28 June	01.28	6.62	0.92-0.08	0.07	4.3 ± 1.9 §
30 June	16.80	0.95	0.48-0.51	0.15	4.4 ± 0.9
8/9 July	15.85	24.82	0.16-0.77	0.38	1.3 ± 1.9
17 July	10.47	2.22	0.32-0.38	0.62	1.2 ± 2.0
29 July	06.37	2.95	0.28-0.35	0.95	10.0 ± 1.7
29/30 July	23.22	1.67	0.69-0.73	0.97	6.6 ± 1.9
05 August	11.93	10.30	0.53-0.78	0.16	6.4 ± 0.9
13 August	05.77	3.65	0.09-0.18	0.37	3.4 ± 1.8
26 August	19.60	2.82	0.07-0.14	0.75	2.8 ± 0.9

* Using the ephemeris of ref. 21, where the centre of eclipse occurs at JD = 2,442,859.7267 with a period of 1.70016779 day.

† Using the 35-day ephemeris derived in the text.

‡ Assuming a power-law spectrum with energy index of -0.48 (from the best fitting ME spectrum) and a low energy cut-off, equivalent to 2×10^{20} H cm $^{-2}$ (ref. 23).

§ For the interval 1983 June 26 06:54 to 07:54 UT (outside of eclipse).

observation¹⁸ on 1984 March 1 detected Her X-1 in a high-state at a flux of ~ 80 mCb, comparable with the previously observed maximum intensity (see also ref. 19).

The interval between the Tenma and EXOSAT observations of high-states in 1983 June and 1984 March, respectively, is approximately 8 times the average 35-day cycle. Figure 1 shows the ME count rates folded using the approximate 35-day ephemeris derived from these observations, which gives X-ray maxima at Julian date (JD) 2,445,478.7 $\pm$ 35.5E. Also given for comparison is the usual envelope of the 35-day X-ray modulation. All the measured intensities were similar to those measured during the first EXOSAT observation. If Her X-1 had been following the established 35-day cycle, we should have detected either a primary or secondary high-state.

The observation on 1983 June 28 was made through an entire eclipse of the X-ray source by its companion and Fig. 2 shows a 2-10-keV light curve from the ME with a hardness ratio obtained from dividing the ME counts between 5 and 10 keV by those between 2 and 5 keV. Coaligned with the ME is a low energy imaging telescope with a channel multiplier array (CMA) in the focal plane; the light curve from this instrument in the 0.02-2.0 keV band²⁰ is given also in Fig. 2. Comparison of the count rates obtained at the start and end of the 1983 June 28 observation with those of the uneclipsed observations suggest that Her X-1 had emerged fully from eclipse by the end of the observation. In contrast to previously observed eclipse transitions^{21,22}, the eclipse ingress and egress are clearly resolved here. At mid-eclipse there is a residual counting rate of ~ 1 count s $^{-1}$ in the ME with no marked change in hardness ratio associated with the eclipse. The CMA data also show residual emission from Her X-1 in the 0.02-2.0-keV band. During a 2.6-h interval centred on mid-eclipse, a count rate of 0.0027 count s $^{-1}$ was detected at a confidence level of $>99\%$. The reduction in intensity in the CMA was ~ 0.25 , similar to that seen in the ME. As the ME array was coaligned for this observation, data from an observation two days later was used to estimate the background subtraction techniques.) Systematic variations in the background counting rate over the period between these two observations could have produced some of the residual flux in the ME at mid-eclipse. However, this is unlikely to be a large effect because a substantial background variation would have caused a change in the hardness ratio, which was not observed. In addition the CMA background is determined simultaneously from a blank-field part of the detector and is not subject to the same uncertainties.

The observation on 1983 August 5 provided the best 2-10 keV X-ray spectrum from the ME. No single continuum model gave an acceptable fit with a power-law model yielding a χ^2 of 105

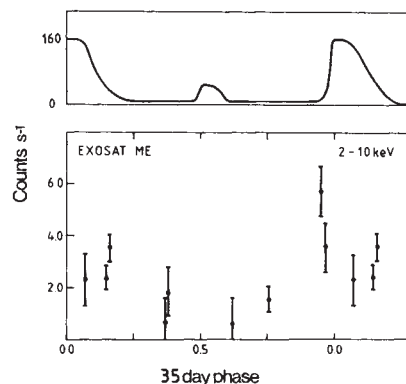


Fig. 1 The lower panel shows the observed ME count rates from Her X-1 averaged over each observation and folded modulo, for the 35-day ephemeris given in the text. The upper panel shows the shape and intensity of the 35-day ME light curve which would have been expected had Her X-1 been behaving normally.

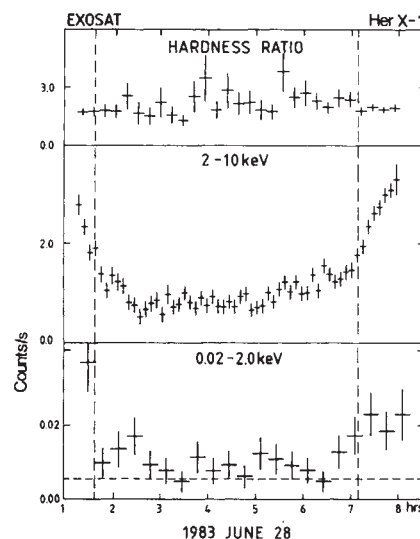


Fig. 2 The ME (2-10 keV) and CMA (0.02-2.0 keV) light curves covering an entire eclipse of Her X-1 with time resolutions of 7 and 20 min, respectively. The dashed vertical lines indicate the expected times of eclipse ingress and egress²¹. The dashed horizontal line indicates the measured CMA background count rate. The hardness ratio is given by the ME counts between 5 and 10 keV divided by those between 2 and 5 keV with a time resolution of 14 min.

for 12 degrees of freedom. The inclusion of an iron K-emission line improved the fit giving a formally acceptable χ^2 of 10 for 9 degrees of freedom for a power law with an energy index of $-0.48^{+0.33}_{-0.08}$ and an equivalent hydrogen column density of $N_H < 1.1 \times 10^{22}$ H cm $^{-2}$. The line strength was $(8 \pm 2) \times 10^{-4}$ ph cm $^{-2}$ s $^{-1}$, corresponding to an equivalent width of 1.2 ± 0.4 keV, with a full-width-half-maximum of < 1.7 keV at an energy of 6.29 ± 0.25 keV (99% confidence limits). The observed spectrum with the best fit model is given in Fig. 3 with the incident spectrum. The 2-10 keV flux level was 6.4×10^{-11} erg cm $^{-2}$ s $^{-1}$ corresponding to a luminosity of 2.6×10^{35} erg s $^{-1}$ at a distance of 5.8 kpc. The spectrum and flux levels are similar to those measured¹⁴ during a previous 35-day low state. About 70% of the < 1 keV high-state flux comes from a 0.11-keV black-body component²³. Extrapolating the above ME spectrum to the CMA energy range, assuming an N_H of 2×10^{20} H cm $^{-2}$, gives an expected CMA count rate of 0.013 ± 0.003 count s $^{-1}$, three times less than the observed count rate of 0.033 ± 0.002 count s $^{-1}$, suggesting that the black-body component was present at an intensity similar to that in the high-state with respect to the high-energy flux.

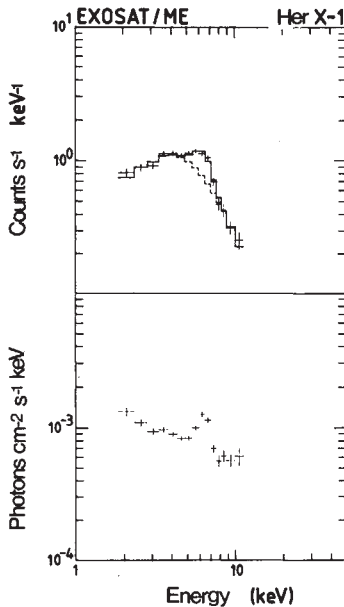


Fig. 3 The upper panel shows the 1983 August 5 ME pulse height spectrum with the best-fitting model shown as a histogram. The lower panel shows the incident spectrum, after deconvolution from the detector response. The dashed line shows the calculated spectrum with the contribution from the iron line excluded.

The duration of the X-ray eclipse ingress or egress of ~ 1 h, when multiplied by the combined X-ray source and secondary velocities of $\sim 300 \text{ km s}^{-1}$, gives a radius for the occulted low-state X-ray emission region of $\sim 5 \times 10^{10} \text{ cm}$, comparable to estimates for the radius of the accretion disk⁸, which confirms the original suggestion¹⁶ that the majority of the low-state X-ray emission comes from X-rays scattered by an accretion disk corona (ADC) from the central source to the observer around the occulting disk. Scattering in an ADC has also been proposed to explain the partial eclipses seen first from 4U1822–37 (ref. 24) and then from 4U2129+47 (ref. 25). There is an asymmetry in the ME ingress and egress light curve, which may be caused by the distortions of the disk responsible for the 35-day cycle. Further observations of eclipses at different 35-day phases could yield further information about a changing disk geometry. The emission seen during mid-eclipse suggests the presence of a second more extended scattering region, with a radius comparable with, or larger than, that of the companion star (HZ Her) ($\sim 2 \times 10^{11} \text{ cm}$), which may be a hot wind excited by X-ray illumination of the face of HZ Her²⁶ and/or the outer regions of the disk²⁷.

Our EXOSAT results indicate that Her X-1 temporarily entered an extended low state presumably caused by a change in the accretion disk that resulted in the disk completely occulting our line of sight to the X-ray pulsar. Jones *et al.*²⁸ (see also ref. 29) examined plates from the archives taken over ~ 80 yr and found that Her X-1 has intervals, lasting typically 10 yr, where the effect of X-ray heating on the companion is not seen suggesting that the X-ray source shuts off completely. Optical and ultraviolet observations of HZ Her between 1983 June and August detected the effects of X-ray heating (ref. 30; J. A. Mattei, I. Howarth, personal communications) and show that Her X-1 had not entered an X-ray inactive state. It is not clear whether the 35-day modulation of the optical light curve was present over the interval after the X-ray cycle had stopped. However, between 1979 and 1982 the amount of X-ray heating seen from HZ Her appears to have decreased by $\sim 15\%$ ³¹. In addition during 1982 and 1983 the measured pulse period was longer than is predicted by extrapolating from previous measurements (refs 13, 32 and T. Ohashi, personal communication). These effects suggest a small reduction in the X-ray luminosity caused presumably by a related change in the mass-transfer rate, which, as the radius of the magnetosphere of the X-ray pulsar is close to the centrifugal radius, could either decrease or increase.

The tilted precessing disk models predict X-ray and optical light curves that are very sensitive to changes in the disk thickness; a small variation in the mass transfer rate could cause the

disk to thicken and completely block our line of sight to the pulsar. In the self-oscillating disk model, a change in the average mass-transfer rate would cause a change in the amplitude of the oscillations in the disk thickness such that the wave minimum would not be deep enough to uncover the X-ray source. As Her X-1 has not entered an X-ray inactive state for 27 yr (ref. 28), further monitoring is essential; the anomalous behaviour in the 35-day cycle may be a precursor to the next X-ray inactive state. With the complications of X-ray heating removed, measurements of the ellipsoidal and radial-velocity variations during the next inactive state will provide additional constraints on the properties of this system.

We thank J. A. Mattei for providing the AAVSO observational data; A. Peacock, B. G. Taylor and the EXOSAT observatory team for their support. A.N.P., N.E.W. and P.B. acknowledge the receipt of ESA Fellowships.

Received 7 August; accepted 24 October 1984.

1. Tannanbaum, H. *et al. Astrophys. J.* **174**, L143–L149 (1972).
2. Giacconi, R. *et al. Astrophys. J.* **184**, 227–236 (1973).
3. Boynton, P. E., Cantera, R., Crosa, L., Deeter, J. & Gerend, D. *Astrophys. J.* **186**, 617–624 (1973).
4. Crosa, L. & Boynton, P. E. *Astrophys. J.* **235**, 999–1015 (1980).
5. Katz, J. I. *Nature* **246**, 87–88 (1973).
6. Roberts, Wm. J. *Astrophys. J.* **187**, 575–584 (1974).
7. Petterson, J. A. *Astrophys. J.* **218**, 783–791 (1977).
8. Gerend, D. & Boynton, P. E. *Astrophys. J.* **209**, 562–573 (1976).
9. Levine, A. M. & Jernigan, J. G. *Astrophys. J.* **262**, 294–300 (1982).
10. Pringle, J. E. in *Accreting Neutron Stars* (eds Brinkman, W. & Trümper, J.) 150–155 (Max-Planck-Institut für Physik und Astrophysik, München, 1983).
11. Meyer, F. & Meyer-Hofmeister, E. *Astr. Astrophys.* **140**, L35–L38 (1984).
12. Taylor, B. G., Andresen, R. D., Peacock, A. & Zobl, R. *Space Sci.* **30**, 479–494 (1981).
13. Nagase, F. *et al. Am. Inst. Phys. Conf. Proc.* (ed. Woosley, S.) **115**, 131–133 (1983).
14. Pravdo, S. H. *et al. Astrophys. J.* **225**, L53–L58 (1978).
15. Parmar, A. N., Sanford, P. W. & Fabian, A. C. *Mon. Not. R. astr. Soc.* **192**, 311–318 (1980).
16. Jones, C. & Forman, W. *Astrophys. J.* **209**, L131–L135 (1976).
17. Turner, M. J. L., Smith, A. & Zimmermann, H. U. *Space Sci.* **30**, 513–524 (1981).
18. Trümper, J. *IAU Circ. No.* 3923 (1984).
19. Giovannelli, F., Kurt, V., Sheffer, E. & Bisnovatij-Kogan, G. S. *IAU Circ. No.* 3924 (1984).
20. De Korte, P. A. J. *et al. Space Sci.* **30**, 495–511 (1981).
21. Deeter, J. E., Boynton, P. E. & Pravdo, S. H. *Astrophys. J.* **247**, 1003–1012 (1981).
22. Davison, P. J. N. & Fabian, A. C. *Mon. Not. R. astr. Soc.* **178**, 27p–29p (1974).
23. McCray, R. A. *et al. Astrophys. J.* **262**, 301–307 (1982).
24. White, N. E. *et al. Astrophys. J.* **247**, 994–1002 (1981).
25. White, N. E. & Holt, S. S. *Astrophys. J.* **257**, 318–337 (1983).
26. McCray, R. & Hatchett, S. *Astrophys. J.* **199**, 196–205 (1975).
27. Begelman, M. C., McKee, C. F. & Shields, G. A. *Astrophys. J.* **271**, 70–88 (1983).
28. Jones, C. A., Forman, W. & Liller, W. *Astrophys. J.* **182**, L109–L112 (1973).
29. Hudec, R. & Wenzel, W. *Bull. astr. Insts. Csl.* **27**, 325–335 (1976).
30. Delgado, A. J., Schmidt, H. U. & Thomas, H. C. *Astr. Astrophys.* **127**, L15–L16 (1983).
31. Thomas, H.-C., Africano, J., Delgado, A. J. & Schmidt, H. U. *Astr. Astrophys.* **126**, 45–50 (1983).
32. Middleditch, J., Pennypacker, C. R. & Burns, M. S. *Astrophys. J.* **274**, 313–326 (1983).

Megaregolith thickness, heat flow, and the bulk composition of the Moon

Kaare L. Rasmussen*† & Paul H. Warren*

* Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California 90024, USA

† Institute of Geophysics, University of Copenhagen, Haraldsgade 6, DK-2200, Copenhagen N, Denmark

The Moon's bulk composition is a major constraint on its origin and evolution. Bulk density alone shows that the Moon is depleted in metallic FeNi relative to the Earth or to chondrites with earthlike MgO/FeO ratios (such as H-group chondrites). Depletions of minor-trace volatile elements are also obvious from geochemical data. The simplest assumption would be that the Moon is not much different in terms of nonsiderophile, nonvolatile elements from the silicate portions of the Earth or H chondrites. Lunar heat flow data¹ have been interpreted to imply that the Moon's uranium content is about 46 ng g^{-1} , about twice that of the Earth's mantle, and three times that of H-chondrite silicates. As it is difficult to envisage any process preceding the formation of planets that would fractionate refractory lithophiles such as uranium from major lithophiles such as Si, disparity in the U/Si ratio implies disparity in provenance or origin. Based on a new model that takes into account the considerable, but variable, thickness of porous, low-

conductivity megaregolith, the thickness of the lunar lithosphere, and the nonrepresentative composition of the crust at one of the two sites where heat flow was measured, we estimate here that the Moon's uranium content is roughly 19 ng g^{-1} . The Moon's bulk composition appears far less exotic than generally assumed.

In the context of planetology, uranium and thorium have importance far out of proportion to their abundance in terms of mass or number of atoms. Other than heat sources associated with the origin of the Moon (for example, the energy of accretion), these two elements are the main sources of lunar internal heat. The Moon's heat production, and consequently its gross evolution, are functions of its U and Th concentrations. The Moon's U and Th contents, inferred from its heat flow, are generally regarded as key constraints on the bulk major element composition of the Moon: cosmochemists consider it unlikely that refractory lithophile elements (Al, Ca, Sc, Ti, V, Sr, Y, Zr, Nb, all 14 rare earth elements, Hf, Ta, Th and U) were fractionated from chondritic (or solar, or 'cosmic') proportions relative to one another at any phase before the aggregation of the Moon out of its protoplanets². Thus, an estimate of the lunar U or Th content can be directly converted into estimates for the major elements Al and Ca. By including estimates of Mg/Fe and Mg/Si based mainly on petrochemical data, cosmochemists estimate bulk-Moon concentrations for numerous elements to three or even four significant figures^{3,4}.

Lunar heat flow was, unfortunately, only measured at two sites (measurements were planned for two others, but Apollo 13 never landed, and the Apollo 16 probe was improperly installed). As revised in 1976¹, the measurements were $21 \pm 3 \text{ erg s}^{-1} \text{ cm}^{-2}$ for the Apollo 15 site and $14 \pm 2 \text{ erg s}^{-1} \text{ cm}^{-2}$ for the Apollo 17 site. Langseth *et al.*¹ argued that these two sites are probably representative for the global mean heat flow, which they estimated to be $18 \text{ erg s}^{-1} \text{ cm}^{-2}$. Assuming this to be steady-state heat flow (current heat production within the Moon = current heat loss from the surface of the Moon), Langseth *et al.* inferred the bulk-Moon U concentration to be roughly 46 ng g^{-1} . They also cautioned that "uncertainty in the global average heat flow cannot be estimated until further data become available", but this caveat has been largely overlooked by cosmochemists. Keihm and Langseth⁵ revised their estimate of global mean heat flow to $14\text{--}18 \text{ erg s}^{-1} \text{ cm}^{-2}$, but did not explicitly revise their estimate of the bulk-Moon U concentration.

We have constructed finite-difference models to assess effects of the Moon's thick low-conductivity megaregolith and convection-free lithosphere layers on its thermal evolution. Previous studies have suggested two potential complications for interpretation of lunar heat flow data. (1) Both sites where heat flow was measured may be unrepresentative⁶, high by as much as a factor of 10, due to the greater thickness of the megaregolith over highlands in comparison to maria, which causes enhanced heat flow over maria, especially near mare-highlands boundaries. The Apollo 15 and 17 sites are both on maria ~5 km from highlands. (2) Heat flow is probably about 1.4 times greater than steady-state models imply, mainly because roughly half of the Moon's mass is conductively cooled lithosphere⁷ (also implicit in the models of Conel and Morton⁶).

Our models are similar to those of Conel and Morton⁶, and we confirm that variations in thickness of megaregolith at mare-highland boundaries probably cause the Apollo 15 and 17 heat flows to be unrepresentative. However, the focusing effect probably only causes heat flow at mare/highland interfaces to be ~15–20% higher than regional averages (Fig. 1). In most models, we assume the U content of the Moon to be 16 ng g^{-1} (that is, 1.0 times non-metallic parts of H-group ordinary chondrites). In all models, U and Th are assumed to be redistributed in the outer 400 km as a result of primordial differentiation: relative to bulk-Moon concentrations, we assume U concentrations to be 4.1 from 0–70 km, 0.56 from 70–130 km, and 0.063 from 130–400 km. Our models are two-dimensional with a spatial resolution of 20 km; they cover 240 km of depth in the Moon, and generally 600 km horizontally. Effects of heat sources below 240 km, as well as residual primordial heat, are simulated by

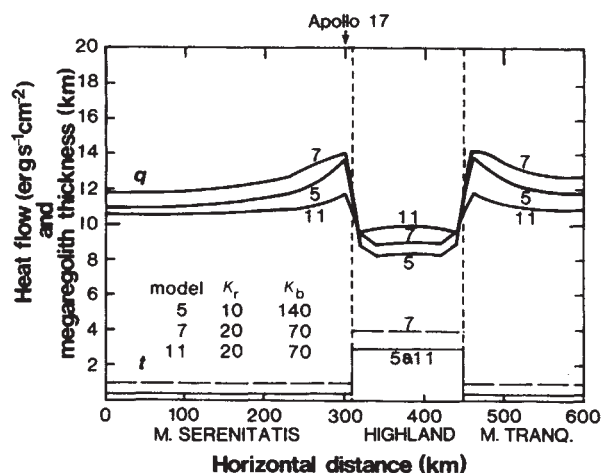


Fig. 1 Typical model results for heat flow (q) at the Apollo 17 site. Also shown are key variables that affect heat flow: t , megaregolith thickness; k_r , thermal conductivity of megaregolith; k_b , thermal conductivity of bedrock. Model results for Apollo 15 site are similar. In all models shown, bulk-Moon U content is 16 ng g^{-1} , redistributed vertically by differentiation.

assuming that the thermal gradient at 240 km is 0.95 K km^{-1} (ref. 6). In conjunction with the assumed bedrock conductivity, this implies a heat flow below 240 km of roughly $2 \text{ erg s}^{-1} \text{ cm}^{-2}$ (half as great as estimated elsewhere¹ for heat flow from below the crust). We find that enhancements of heat flow at mare/highlands boundaries (Fig. 1) are less than originally suggested by Conel and Morton⁶, because our models involve more realistic assumptions about lateral variations in megaregolith thickness, and about the difference in thermal conductivity of megaregolith as opposed to bedrock. Also, our models take into account the finite dimensions of the highland regions near the two landing sites in question.

In previous models⁶, the ratio (highlands megaregolith thickness/mare megaregolith thickness) was varied from 10 to 1,000. Numerous lines of evidence^{8–11} indicate that the megaregolith is generally 1–3 km thick under highlands. Highlands cover about 83% of the total surface; mare basalts cover 17%. Formation of impact basins tends to cause ejection of uppermost crustal material, including former megaregolith, radially outward from the point of impact. At the centres of comparatively recent impact basins such as Imbrium (adjacent to the Apollo 15 site) and Serenitatis (adjacent to the Apollo 17 site) there is presumably no appreciable megaregolith. Just outside of recent impact basins megaregolith thickness is probably above average, ~3 km. Under maria within impact basins, whatever megaregolith that was not stripped away during the basin-forming impact must occur as layer(s) sandwiched between nonmare bedrock below, and layers of mare basalt frequently >1 km in total thickness¹² above. Graben geometries⁹ suggest that the crust becomes more coherent (megaregolith gives way to bedrock) at a depth generally 1.2–1.9 km below the surface under maria, as opposed to mostly 2–3 km below the surface under highlands. Analysis of 'ghost' craters¹³ as well as radar sounding data¹⁴ suggest that the basalt layer under Mare Serenitatis is >2 km thick within 110 km of the Apollo 17 site; so unless the depth to nonmare bedrock is exceptionally great, megaregolith thickness under Mare Serenitatis adjacent to the Apollo 17 site is probably <0.5 km. Our estimates for megaregolith thicknesses close to the Apollo 17 site are 0.2–0.6 km for the mare and 2–3 km for the highlands. The Apollo 15 site is well beyond the original rim of the Imbrium basin¹⁵ and the adjacent portion of Mare Imbrium is not particularly deep¹⁶. The elevations reached by the adjacent highlands suggest an exceptionally thick basin-ejecta deposit, however. The Apennine Mountains 10 km from the landing site are nearly 4 km higher in elevation. Our estimates for megaregolith thicknesses close to the Apollo 15 site are 0.5–1.5 km for the mare and 3–4 km for the highlands. The ratio

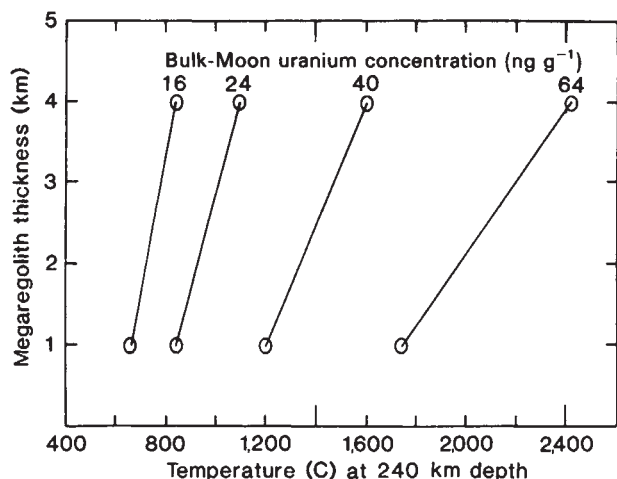


Fig. 2 Present temperature at 240 km depth as a function of bulk-Moon U content and megaregolith thickness. Calculated assuming conservatively high conductivities (for megaregolith, $k_r = 40 \text{ kerg s}^{-1} \text{ cm}^{-1} \text{ K}^{-1}$; for bedrock, $k_b = 200 \text{ kerg s}^{-1} \text{ cm}^{-1} \text{ K}^{-1}$).

(highlands megaregolith thickness/mare megaregolith thickness) is marginally lower for the Apollo 15 site than for the Apollo 17 site.

In previous models⁶, the ratio (bedrock conductivity/megaregolith conductivity) varied from 10 to 1,000. Among lunar materials, there is a fairly good linear correlation between porosity and the logarithm of thermal conductivity¹. The megaregolith consists of porous impact breccias, whose porosities, if not directly measured, can be inferred from measured densities. The mean porosity of 39 lunar breccias, including eight impact melt breccias, is $17 \pm 10\%$ (refs 17, 18). The mean porosity of Apollo 15 and 17 mare basalts, about 10% (ref. 19), is potentially misleading. When these rocks were excavated by impacts, the vesicular upper layers of the lava flows were probably systematically oversampled. A second-order effect comes from compositional variations: mafic silicates have 2–3 times higher conductivity than plagioclase²⁰. Based on imprecise porosity constraints, the conductivity–porosity relationship¹ indicates that megaregolith conductivity is roughly (within a factor of 3) $20 \text{ kerg s}^{-1} \text{ cm}^{-1} \text{ K}^{-1}$, and bedrock conductivity is roughly (within a factor of 2) $150 \text{ kerg s}^{-1} \text{ cm}^{-1} \text{ K}^{-1}$.

Besides focusing heat flow at mare/highland boundaries (Fig. 1), the megaregolith reinforces the effect of a thick lithosphere⁷, making for a higher than steady-state present heat flow, and an interior that is hotter than would otherwise be the case. We have examined the effect of varying megaregolith thickness by running models similar to those illustrated in Fig. 1, except that megaregolith thickness was assumed to be laterally uniform. Increasing the megaregolith thickness in these models from 0.1 to 2 km resulted in an 11% increase in present heat flow, and a 112 °C increase in present temperature at 240 km depth. Magnetometer data indicate that the temperature at 240 km is in the range 300–900 °C, and most probably in the range 500–700 °C (ref. 21). An absolute upper limit for the 240 km temperature is the 12-kbar solidus of peridotite, which is $\sim 1,300$ °C (ref. 22). The highlands (83% of the Moon's surface) are covered by 1–3 km of megaregolith. Figure 2 shows the 240-km temperature predicted by our model calculations as a function of megaregolith thickness and heat production (bulk-Moon U content). These calculations indicate that if the 240 km temperature is < 900 °C and megaregolith thickness averages even 1 km, then the bulk-Moon U content must be $< 30 \text{ ng g}^{-1}$. Even 20 ng g^{-1} of U is only barely consistent with the latest estimate of the actual temperature (400–700 °C) (ref. 21).

One of the two measured heat flow sites, Apollo 15, is probably less representative because it is in an exceptionally U-rich region

of the Moon. Measured heat flow is 50% greater for the Apollo 15 site than for the Apollo 17 site¹. We find it unlikely that heat flow focusing is significantly more pronounced at Apollo 15 than at Apollo 17; the reverse is more likely, because the ratio (highlands megaregolith thickness/mare megaregolith thickness) is probably lower in the Apollo 15 area than in the Apollo 17 area. Orbital spectrometry data^{23,24} show that in most of the central-western nearside, regolith U and Th contents are far above the global means. In the vicinity (a circle roughly 80 km in radius) of the Apollo 15 site, U and Th contents are 3.3 times the global means; in the vicinity of the Apollo 17 site the U and Th contents are lower, but still 1.4 times the global means²³. It might be argued that these regional enrichments are purely surficial. If the central-western nearside surface concentrations were representative of the entire vertical range of the crust, the lower crust would be partially molten¹. We agree with Langseth *et al.*¹ that the upper crust is probably enriched in U and Th. Most of the U and Th in the crust probably occurs as KREEP²⁵, basaltic material named for its high contents of incompatible elements such as K, rare earth elements and P. Most pristine nonmare lithologies appear to have formed as intrusive cumulates, but all of the KREEPy lithologies have volcanic or hypabyssal textures²⁵, which implies that KREEP is probably enriched in the shallowest crust. Nevertheless, a compositional anomaly over a region the size of the central-western nearside probably reflects regional compositional variations within a major fraction of the vertical range of the crust²⁶, not just the upper few kilometres. If KREEP originated as the residual liquid of the primordial 'magma ocean', it probably was concentrated under regions with relatively thin crust²⁵. Gravity data imply that the central-western nearside is such a region²⁷. This interpretation is borne out by the marked tendency for pristine anorthositic and magnesian-troctolitic intrusive cumulates from Apollo 14 (the most KREEPy landing site) to have higher incompatible element contents than their counterparts from other sites²⁸. We conclude that only the Apollo 17 site, and not the Apollo 15 site, is likely to be representative of the global mean for lunar heat flow.

Correcting the Apollo 17 measurement for the effect (enhancement by 15–20%) of locally thin megaregolith, our best estimate for present global mean heat flow is $12 \text{ erg s}^{-1} \text{ cm}^{-2}$, which is only 0.67 times that estimated in connection with the currently accepted estimate for the bulk-Moon U content¹. Note that the Langseth *et al.*¹ estimate for the bulk-Moon U content was based on the assumption that heat flow is steady-state. The Moon's thick lithosphere probably causes present heat flow to be about 1.4 times present heat production⁷. Our calculations indicate an additional enhancement of 1.1–1.2 times steady-state heat flow, as a result of heat accumulation caused by the insulating effect of the generally 1–2 km thick megaregolith. Assuming that the Moon has a chondritic Th/U ratio, steady-state heat flow is $0.39 \text{ erg s}^{-1} \text{ cm}^{-2}$ for every ng g^{-1} of U in the Moon. Thus, if present heat flow is 1.5 times steady-state, a global mean heat flow of $11 \text{ erg s}^{-1} \text{ cm}^{-2}$ implies a bulk-Moon U content of 19 ng g^{-1} , which is only 0.41 times that estimated by Langseth *et al.*¹.

Accepting the common assumption that bulk-Moon refractory lithophile elements are unfractionated relative to chondritic proportions², a lower bulk-Moon U content has important implications for the major element bulk composition of the Moon, because Al and Ca are also refractory lithophiles. In the early 1970s, estimates of the bulk-Moon Al_2O_3 content were as high as 17 or even 27 wt% (refs 29, 30). Chondritic proportionality to 19 ng g^{-1} of U implies an Al_2O_3 content of 3.2 wt%. Lower U and Al_2O_3 contents have important implications for models of lunar evolution. For example, one of the main arguments in favour of a deep primordial magma ocean is that the crust appears to be greatly enriched in Al_2O_3 relative to the bulk Moon³. Accepting our lower bulk-Moon Al_2O_3 content, the implied crustal Al_2O_3 enrichment becomes even more extreme. Lower bulk-Moon U and Al_2O_3 contents also have important implications for the origin of the Moon. Aside from depletions

of FeNi and volatile elements, the Moon's bulk composition may be less exotic than generally assumed. Non-metallic parts of H-group ordinary chondrites contain about $16 \text{ ng g}^{-1} \text{ U}$ and $2.7 \text{ wt\% Al}_2\text{O}_3$ (ref. 31). According to Ringwood²², the Earth's upper mantle contains about $21 \text{ ng g}^{-1} \text{ U}$ and $3.3 \text{ wt\% Al}_2\text{O}_3$. There is no reason based on lunar heat flow to rule out a close genetic relationship between the Moon and either of these materials.

We thank J. T. Wasson for helpful discussions and M. G. Langseth for a constructive review. This research was supported by NASA grant NAG 9-87 and by the Danish Natural Science Research Council.

Received 3 September; accepted 29 October 1984.

- Langseth, M. G., Jr., Keihm, S. J. & Peters, K. *Proc. 7th Lunar Sci. Conf.* 3143-3171 (1976).
- Warren, P. H. in *Workshop on Pristine Highlands Rocks and the Early History of the Moon* (eds Longhi, J. & Ryder, G.) (Lunar and Planetary Institute Technical Rep. 83-02, 75-79, 1983); *Geochim. cosmochim. Acta* 47, 1559-1571 (1983).
- Taylor, S. R. *Planetary Science: A Lunar Perspective* (Lunar and Planetary Institute, Houston, 1982).
- Morgan, J. W., Hertogen, J. & Anders, E. *Moon and Planets* 18, 465-478 (1978).
- Keihm, S. J. & Langseth, M. G. *Proc. 8th Lunar Sci. Conf.* 499-514 (1977).
- Conel, J. E. & Morton, J. B. *The Moon* 14, 263-289 (1975).
- Schubert, G., Stevenson, D. & Cassen, P. *J. geophys. Res.* 85, 2531-2538 (1980).
- Aggarwal, H. R. & Oberbeck, V. R. *Proc. 10th Lunar planet. Sci. Conf.* 2689-2705 (1979).
- Golombek, M. P. *J. geophys. Res.* 84, 4657-4666 (1979).
- Thompson, T. W. *et al. Moon Planets* 21, 319-342 (1979).
- Head, J. W. *Proc. 7th Lunar Sci. Conf.* 2913-2929 (1976).
- Horz, F. *Proc. 9th Lunar planet. Sci. Conf.* 3311-3331 (1978).
- DeHon, R. A. & Waskom, J. D. *Proc. 7th Lunar Sci. Conf.* 2729-2746 (1976).
- Peebles, W. J. *et al. J. geophys. Res.* 83, 3459-3468 (1978).
- Head, J. W. & Settle, M. *Imbrium Consortium Vol. 1* (ed. Wood, J. A.) 5-14 (Smithsonian Astrophysics Observatory, Massachusetts, 1976).
- DeHon, R. A. *Proc. 10th Lunar planet. Sci. Conf.* 2935-2955 (1979).
- Talwani, M., Thompson, G., Dent, B., Kahle, H.-G. & Buck, S. *Apollo 17 Preliminary Science Rep.* (ed. Parker, R. A.) 13.1-13.13 (NASA, Washington, 1973).
- Horai, K. & Winkler, J. L. Jr *Proc. 7th Lunar Sci. Conf.* 3183-3204 (1976).
- Apollo Lunar Sample Information Catalogs* (Lunar Sample Curator, NASA Johnson Space Center, Houston).
- Horai, K. *J. geophys. Res.* 76, 1278-1308 (1971).
- Hood, L. L., Herbert, F. & Sonett, C. P. *J. geophys. Res.* 87, A109-A116 (1982).
- Ringwood, A. E. *Origin of the Earth and Moon* (Springer, New York, 1979).
- Metzger, A. E., Haines, E. L., Parker, R. E. & Radocinski, R. G. *Proc. 8th Lunar Sci. Conf.* 949-999 (1977).
- Surkov, Yu. A. *Proc. 12th Lunar planet. Sci. Conf.* 1377-1386 (1981).
- Warren, P. H. & Wasson, J. T. *Rev. Geophys. Space Phys.* 17, 73-88 (1979).
- Binder, A. B. *The Moon* 14, 237-245 (1975).
- Bills, B. G. & Ferri, A. J. *J. geophys. Res.* 82, 1306-1314 (1977).
- Warren, P. H. *et al. J. geophys. Res.* 88, B151-B164 (1983).
- Wanke, H. *et al. Proc. 5th Lunar Sci. Conf.* 1307-1335 (1974).
- Anderson, D. L. *Earth planet. Sci. Lett.* 18, 301-316 (1973).
- Mason, B. *US geol. Surv. Prof. Pap.* 440-B, Part 1 (1979).

Evidence of atmospheric gravity waves produced during the 11 June 1983 total solar eclipse

E. J. Seykora*, A. Bhatnagar†, R. M. Jain† & J. L. Streeter‡

* East Carolina University, Greenville, North Carolina 27834, USA

† Udaipur Solar Observatory, Udaipur, India

‡ Southwestern at Memphis, Memphis, Tennessee 38112, USA

During a solar eclipse the Moon's shadow moves at supersonic speed through the Earth's atmosphere. Chimonas and Hines^{1,2} suggested that the resultant cooling of the atmosphere would generate a bow wave of atmospheric gravity waves, which may be detectable as a travelling ionospheric disturbance or as a ground-level atmospheric pressure variation. Although the evidence for gravity waves in the ionosphere from a solar eclipse is still weak, they may have been detected³. Ground-level pressure changes during solar eclipses occur near the region of totality⁴, but the measured wave velocity and period do not agree with that of the predicted disturbance^{5,6}. We now report the detection of a ground-level pressure wave detected at three stations in India and one station in Java, Indonesia. These data may provide the first direct observation of eclipse generated gravity waves over a very long range. The most distant station in India was 6,600 km from the eclipse centre line. The microbarometer recordings indicate that a wave disturbance was recorded at each station with a quasi-period of ~4 h and a wave velocity of ~320 m s⁻¹.

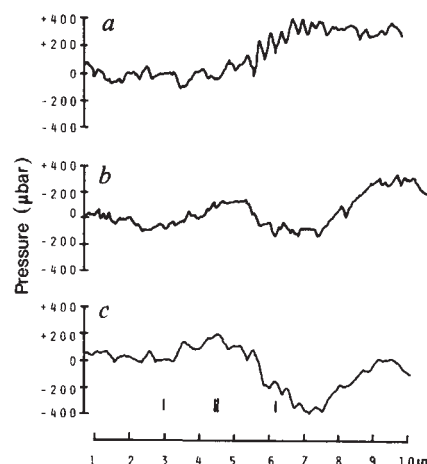


Fig. 1 Ground-level atmospheric pressure fluctuations at eclipse centre line East Java: a, 9 June; b, eclipse day, 11 June; c, residual fluctuations, after removing daily tides, during the 11 June 1983 solar eclipse. First to fourth contact times are indicated.

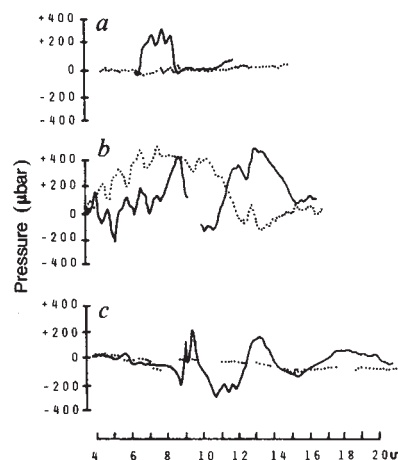


Fig. 2 Ground-level atmospheric pressure fluctuations in India during the solar eclipse and control days: a, Trivandrum, 10 June (·····); 11 June (—); b, Udaipur, 10 June (·····); 11 June (—); c, Leh, 11 June (—); 12 June (·····). The altitude at Leh was 10,800 ft, all other stations were at sea level.

All four recording stations used identical microbarometers, operated in the differential mode and adjusted to respond to pressure variations from 1 s for short-term variations to >5 h for long-term variations. The microbarometer stations in India were arranged on a line approximately perpendicular to the eclipse path, which was 3,716 km south of India. They were located at Trivandrum, near the south tip of India, Udaipur in central India, and Leh, northern India ~6,600 km from the line of totality. The recording stations in India were outside the penumbra region of the eclipse, and represent far field stations in that the solar insolation at these stations was not affected by the eclipse. A fourth recording station was set up on the centre line of totality in East Java, Indonesia. The microbarometer recordings on the eclipse day and on several previous days are unique recordings in that the atmosphere was quite stable at each of the recording stations. The recordings from each station will be discussed in turn, followed by a general description of the atmospheric wave motion which resulted from the eclipse. Centre line East Java: Figure 1a shows one of several microbarometer recordings taken before the day of the eclipse. The upward trend in pressure results from thermal-barometric tides in this area and exhibited very little variation from day to day. The oscillation recorded between 05.00 and 08.00 UT represents atmospheric gravity waves which occur from time to time and typically have periods of ~20 min. Figure 1b represents the microbarometer recording on 11 June, which indicates the pres-

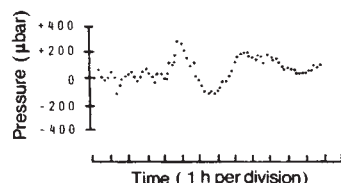


Fig. 3 Three-station average of the ground-level atmospheric pressure fluctuation in India on 11 June 1983. The pressure profiles were normalized to the same arrival time, based on a wave velocity of 320 m s^{-1} , before averaging.

ence of very low frequency components compared with the recordings on background days. After subtraction of the thermal tides from Fig. 1b, Fig. 1c shows the eclipse-induced pressure variations at the centre line of the eclipse. The predominant period of the oscillation at this site was $\sim 4\text{--}4.5 \text{ h}$, and the resultant wave motion continued $\sim 7 \text{ h}$ after first contact.

Trivandrum, India: This station, some 3,700 km from the centre line, had very low background pressure fluctuations and very little atmospheric tidal variation. Unfortunately, commercial power was lost at this station on 11 June but was restored using an auxiliary generator between 06.30 and 11.30 UT. The expected arrival time for the pressure disturbance at Trivandrum was between 06.00 and 06.30 UT. Figure 2a shows the eclipse day recording corrected for background. As shown in Fig. 2a, a rather large pressure wave or pulse was recorded between 06.30–08.30 UT. Furthermore, the pulse or wave contained rather low frequency components. On no other day was a similar pressure wave recorded at this station. For comparison, the pressure fluctuations on 10 June are also shown in Fig. 2a.

Udaipur, India: At Udaipur, 5,500 km from the centre line, typical daily tidal and random pressure fluctuations were rather large on 11 June, and on all previous days that measurements were taken. Figure 2b represents the difference between the eclipse day and a typical background day. The resultant recording shows a large high-frequency fluctuation because of local atmospheric turbulence, although between 08.00 and 15.00 UT a low frequency pressure variation is evident with a quasi-period of 4.5 h. The pressure fluctuations on 10 June are shown in Fig. 2b.

Leh, India: Figure 1c shows the microbarometer recording for 11 June at Leh, 6,600 km from the centre line. The background days recorded for 10 June and 12 June were interrupted by many power failures. During the intervals of operation on those days, the atmospheric pressure was stable, showing no unusual variations of pressure. The pressure fluctuations on 12 June are also shown in Fig. 2c. A rather large low frequency wave was recorded 11 June on an otherwise very stable day. As Fig. 2c shows, the predominant oscillation period was 4 h with the onset of the oscillation at approximately 09.00 UT.

Although each of the pressure profiles in India was deformed by local conditions and random pressure fluctuations, the time delay of the pressure distribution at each station is important. The delay between Trivandrum and Udaipur was $\sim 1.5 \text{ h}$ and that between Udaipur and Leh, 1 h. The resultant wave velocity would then be $\sim 330 \text{ m s}^{-1}$ between Trivandrum and Udaipur, and 306 m s^{-1} between Udaipur and Leh. The wave velocity as calculated from the centre line to Trivandrum corresponds to $\sim 300 \text{ m s}^{-1}$.

Assuming a wave velocity of $\sim 320 \text{ m s}^{-1}$, the pressure profiles were shifted such that each had the same arrival time, and each of these curves was averaged for the times where data were available. This average pressure profile for the three stations in India is shown in Fig. 3a. In this averaging process, we have neglected any effects of dispersion which may occur in distances less than one wavelength, during the transit of the pressure disturbance from Trivandrum to Leh. As expected, and shown in Fig. 3, the higher frequency random fluctuations are reduced and the remaining disturbance reveals a wave disturbance whose characteristics seem to be consistent with: wave velocity, $\sim 320 \text{ m s}^{-1}$; quasi-period, $\sim 4 \text{ h}$; amplitude, $\sim 100\text{--}150 \mu\text{bar}$.

The corresponding air temperature for a speed of sound of 320 m s^{-1} is 253 K. This temperature also corresponds to an altitude of 50 km, or in the ozonosphere. These recorded data are consistent with the wake of atmospheric gravity waves predicted by Chimonas and Hines for solar eclipses. Their predicted wave period was $\sim 4 \text{ h}$ with a wave velocity of $\sim 300 \text{ m s}^{-1}$. These data may represent the first detection of this extremely low frequency disturbance over a very long range.

We thank Dr M. Aizenman, the NSF Eclipse Coordinator, and Mrs K. Crouch, expedition coordinator, for their support. This research is based on work supported by the NSF. Part of the research work carried out in India was supported by the Indian Space Research Organisation.

Received 16 August; accepted 30 October 1984.

1. Chimonas, G. & Hines, C. O. *J. geophys. Res.* **75**, 875 (1970).
2. Chimonas, G. *J. geophys. Res.* **75**, 5545–5551 (1970).
3. Davis, M. J. & da Rosa, V. A. *Nature* **226**, 1123 (1970).
4. Anderson, R. C., Keeter, D. R. & Myers, O. E. *J. Atmos. Sci.* **29**, 583–587 (1972).
5. Davies, K. *Proc. Ind. Nat. Sci. Acad.* **48A**, Suppl. No. 3, 342–355 (1982).
6. Reddy, C. A. *Proc. Ind. Nat. Sci. Acad.* **48A**, Suppl. No. 3, 356–369 (1982).

Thinning of the lithosphere by small-scale convective destabilization

David A. Yuen* & Luce Fleitout†

* Department of Geology, Arizona State University, Tempe, Arizona 85287, USA

† Laboratoire de Géophysique et Géodynamique interne, Université de Paris-Sud, Bat. 510, 91404 Orsay, France

Gravitational, topographical and seismological studies of uplift regions around hot spots on oceanic,^{1,2} and continental plates^{3–7} have suggested that considerable lithospheric thinning has taken place beneath these areas. Previous models used to study this phenomenon have been based on thermal conduction^{8–11} and on the dynamic interaction between large-scale plume structures¹² and the lithosphere, for a purely temperature-dependent viscosity. However, their predictions of the time scale for direct erosion¹² of the lithosphere by these upwellings are too long. We propose here a new dynamical mechanism which works most efficiently for a temperature and pressure-dependent rheology. Our idea is that thinning of the lithosphere is accomplished by strong secondary convection enhanced by the presence of a low-viscosity zone which grows from a plume associated with a hot spot. This process is capable of producing the rapid time scale of the swell uplift¹³, of the order of 10 Myr, and the rates of thinning, of order 10 km Myr^{-1} , required by both geophysical^{13,14} and geochemical¹⁵ observations.

Broad highlands¹⁶ in several geological areas, such as the Hoggar Massif, the East African plateaus and the northern Rocky Mountains surrounding Yellowstone, can be explained by a rapid increase in the lithospheric temperatures on a time scale of only $\sim 10^7 \text{ yr}$. Similar swells can also be found in oceanic areas, for example, over Hawaii¹¹, which are characterized by a surface elevation of 2–3 km and a horizontal scale of 10^3 km . The geoid height anomalies over the Bermuda Rise² and Hawaiian Swell¹ are consistent with a shallow depth of compensation, indicating a compensating mass deficiency at 40–60-km depth. Anomalously low P_n velocity and thinned crust^{6,7}, as detected by seismic refraction, have often been interpreted as evidence of an asthenospheric diapir in contact with the base of the crust at the Moho. Furthermore, heat transport into the lower lithosphere has been invoked to account for its differential stretching during the formation of some sedimentary basins¹⁷. All the geophysical evidence seems to support the idea that lithospheric thinning, by means of additional heat transport, can occur beneath hot spots.

Various mechanisms have been proposed to account for rapid lithospheric heating. Uplift and magmatism on the Colorado

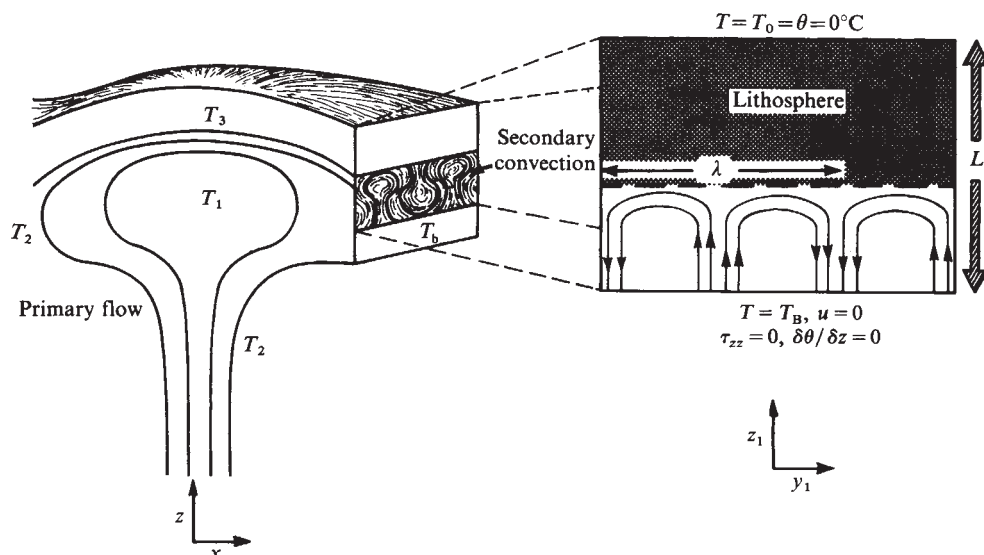


Fig. 1 Schematic diagram portraying the development of secondary convection from a rising primary plume. The projection (right) describes the mathematical idealization of this process by the mean-field equations. Isotherms are depicted such that $T_1 > T_b > T_2 > T_3$. The intrusion temperature T_b is highest in the centre of the plume and decreases with x . The y_1 -axis associated with the secondary roll may not be orthogonal to the x_1 -axis. The lithosphere (dotted background) shrinks with time owing to enhanced heat transfer. An upwardly migrating low-viscosity zone is formed right below the lithosphere, owing to the very nature of $\eta(T, p)$.

plateau have been explained by a delamination event¹⁸ involving the detachment of the lower lithosphere from the crust owing to a low-viscosity layer in the lower crust. Diapiric penetration of the lithosphere caused by Rayleigh–Taylor instabilities has been suggested for the Rio Grande Rift and the Rhenish Massif^{10,19,20}. Another mechanism has involved a conduction model⁸ with a moving asthenosphere–lithosphere boundary on which an arbitrarily high heat flux was imposed. In yet another work¹² the ability of a hot plume to supply a large heat input was studied in the context of a boundary layer flow model in which temperature-dependent viscosity, but not buoyancy forces, was included. However, this type of stagnation flow proved unable, over a wide range of temperatures, to thin the lithosphere sufficiently fast.

Here we propose a dynamical model combining both convective thinning and destabilization of the lithosphere owing to its own buoyancy forces. In this process the material at the lithosphere–asthenosphere boundary is warmed conductively by the extra heat transported by secondary convection in the asthenosphere. This matter consequently becomes soft enough to creep but, being denser than the underlying soft asthenosphere, it will plunge into it as a blob, thereby efficiently eroding the lithosphere.

Our mechanism is predicated on the belief that strong small-scale convection can develop in the low-viscosity zone from a primary flow, corresponding to a steady rising plume found in numerical calculations^{21,22} and laboratory experiments^{23,24}. This zone arises because of the dependency of the mantle viscosity on the combined effects of temperature and pressure. Previous notions^{25,26} of small-scale convection, as a means of producing lower lithospheric instability, were based on the assumption of a low-viscosity zone within the framework of constant viscosity. This type of rheology, however, does not allow for the lithosphere to be thinned dynamically. A continuous, voluminous supply of hot material is thus assumed in order to maintain this primary flow. For high Rayleigh number convection, characteristic of the mantle, the mass flux associated with upwelling plumes should be able to sweep away the cold material falling from the destabilized lithosphere. Fast plumes then prevent the interior from cooling and thus help to maintain an isothermal interior at T_b . This scenario is depicted in Fig. 1. However, these instabilities are unlikely to be manifested in the laboratory experiments because the pressure-dependence of the viscosity cannot be simulated at all in the laboratory. We may regard presently the orientations of these secondary rolls as being random and, not in the strictest sense, as being cross rolls with a longitudinal direction²⁷. Thus the y_1 -axis may not be orthogonal to the x_1 -axis in Fig. 1. Clearly, future three dimensional calculations must investigate in greater detail the structure of this new type of secondary motion.

We will construct here a simple mathematical model to simulate the secondary convection. The mathematical idealization of the physical situation is shown in the projection (on the right) of Fig. 1. As computation of this flow is quite difficult to handle using the full set of equations, we have employed an approximate procedure, called the mean-field equation²⁸, to study the non-linear evolution of secondary convection caused by the rise of hot material into the asthenosphere. Briefly, the mean-field approximation consists of describing the horizontal dependence of the flow field by a single periodic mode and then solving the resultant non-linear partial differential equations in time and the vertical coordinate, z . The depth of this layer, L , is taken to be 300 km. The creep law, as described above, depends on both temperature and pressure because studies²⁹ of the thermal state beneath the continents have shown that $\eta(T, p)$ allows more heat to be transported than does a purely temperature-dependent rheology for similar values of the interior viscosity.

The set of single-mode mean-field equations^{28,29} describing the time-dependence of the mean-temperature, T , the temperature fluctuation, θ , and the vertical velocity, w , for a variable viscosity medium²⁹ are given in dimensionless form by

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial z^2} - \frac{\partial}{\partial z}(w\theta) \quad (1)$$

$$\frac{\partial \theta}{\partial t} = \frac{\partial^2 \theta}{\partial z^2} - k^2 \theta - w \frac{\partial T}{\partial z} \quad (2)$$

$$\begin{aligned} \frac{\partial^4}{\partial z^4} - 2k^2 \frac{\partial^2}{\partial z^2} + k^4 + \frac{2}{\eta} \frac{\partial \eta}{\partial z} \left(\frac{\partial^3}{\partial z^3} - k^2 \frac{\partial}{\partial z} \right) \\ + \frac{1}{\eta} \frac{\partial^2 \eta}{\partial z^2} \left(\frac{\partial^2}{\partial z^2} + k^2 \right) w + \frac{Rk^2 \theta}{\eta} = 0 \end{aligned} \quad (3)$$

The dimensionless wavenumber, k , is denoted by $2\pi L/\lambda$, where λ is the wavelength of the horizontal mode under consideration. The reference Rayleigh number, R , is employed here solely for the purpose of rendering the equations dimensionless and is given by

$$R = \frac{\alpha g \Delta T L^3}{\kappa \nu_r} \quad (4)$$

where $\alpha = 3 \times 10^{-5} \text{ K}^{-1}$ is the thermal expansion coefficient, $\kappa = 0.008 \text{ cm}^2 \text{ s}^{-1}$ is the thermal diffusivity, g is the gravitational acceleration, $\Delta T = T_b - T_0$ is the temperature drop and ν_r is a reference viscosity. The rheological parameters used throughout have an activation energy of 100 kcal mol^{-1} , an activation volume of 12 $\text{cm}^3 \text{ mol}^{-1}$, and a pre-exponential parameter such that the viscosity at a depth of 300 km and a temperature of 1,300 °C takes on a value of 10^{22} P . These values are selected because they would best exemplify the subsolidus deformation behaviour of mantle materials.

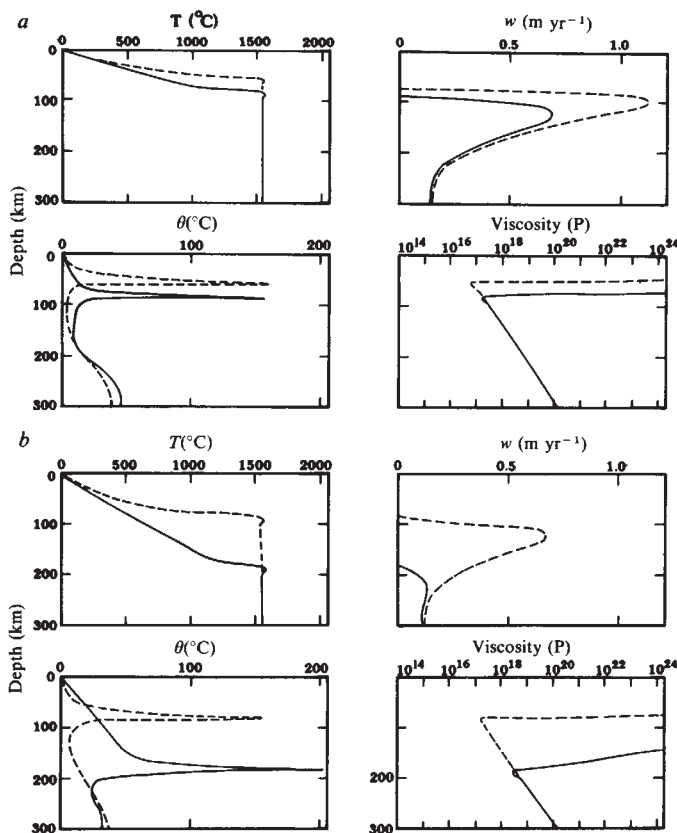


Fig. 2 Mean-temperature, vertical velocity, temperature perturbation and the mean viscosity for lithospheric thinning. The initial thicknesses of the lithosphere are: *a*, 100 km and *b*, 200 km. The basal temperature T_b is 1,540 °C and the wavelength is 150 km, around which maximum heat transfer is obtained. The activation energy is 100 kcal mol⁻¹ and the activation volume is 12 cm³ mol⁻¹. Solid curves represent $t = 3.1$ Myr, dashed curves stand for $t = 14.1$ yr (*a*) and 51.8 Myr (*b*).

The numerical methods for integrating equations (1)–(3) are described elsewhere²⁹. As shown in Fig. 1, the lowest boundary conditions are applied at $z = L$ above which most of the return flow of the secondary rolls takes place. (Note that the horizontal velocity is $\partial w / \partial z$ in Fig. 2.) Both the horizontal velocity, u , and the normal stress, τ_{zz} , are imposed to vanish at this boundary. There is mass flowing across $z = L$. Below the region of vigorous small-scale convection, the fluctuating temperature becomes flat; hence we may set $\partial \theta / \partial z = 0$ there. We have found that changing the mechanical boundary conditions at $z = L$ does not influence the overall results, as long as there is an upward mass flux across $z = L$.

The initial temperature is a linearly-conducting profile through a lithosphere with a gradient of 6.5 °C km⁻¹. Two different initial lithospheric thicknesses have been examined: the first, with a thickness of 100 km, is meant to describe an oceanic lithosphere; the second, with a thickness of 200 km, should characterize an old continental lithosphere. The initial temperature below the lithosphere assumed a constant value T_b , a varying parameter, which must be greater than T_E , the equilibrium temperature of the upper mantle associated with a lithosphere thickness given for the above rheological parameters. We may then regard $T_b - T_E$ as a finite-amplitude perturbation of the basic convecting state owing to the upward penetration of a hot plume. An increase in the temperature of the asthenosphere generates vigorous secondary motion, leading to thinning of the lithosphere by enhanced convective heat transfer.

Figure 2 illustrates the basic physics of this mechanism, where T_b is set to 1,540 °C, corresponding to an increase of 240 °C above the ambient values of T_E . One observes that there is a sharp change in the temperature gradient at the base of the

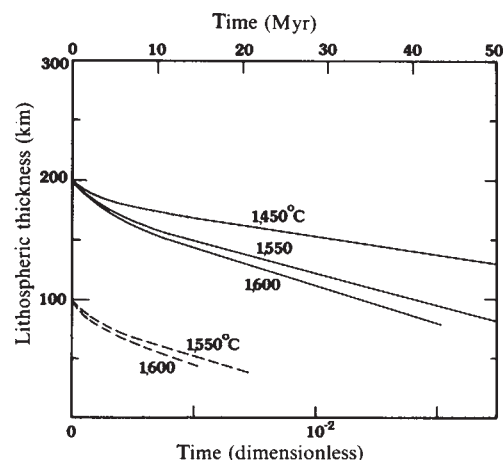


Fig. 3 Time history of lithospheric thinning for various values of T_b . Solid and dashed lines represent, respectively, initial plate thicknesses of 200 and 100 km. The rheological parameters are the same as those in Fig. 2. Temperatures adjacent to the lines represent the intrusion temperature T_b . The wavelength is 150 km.

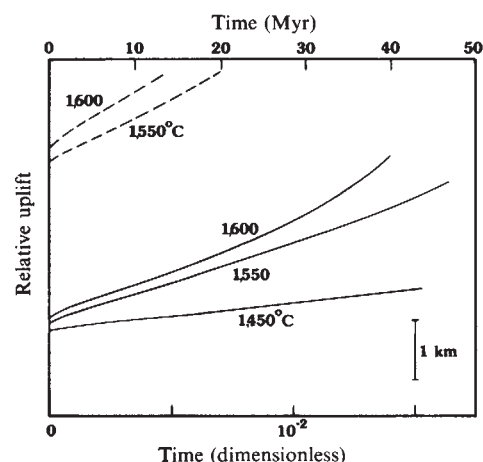


Fig. 4 The temporal evolution of the relative elevation. The temperature has been averaged over the top 200 km. Otherwise, the same symbols and parameter values are used as in Fig. 3.

lithosphere T indicating a large heat input into the plate and resulting in lithospheric thinning. The place where T exhibits a sharp kink migrates upward at a rate of about 2 km Myr⁻¹ for both *a* and *b*. The vertical velocity increases dramatically with time assuming a jet-like profile in a low-velocity zone with a minimum of $O(10^{17}$ P). The narrowness of the low-viscosity zone would be difficult to resolve from loading responses. The vertical velocity is faster in the case of a thinner initial lithosphere, as the viscosity minimum is lower and is able to migrate rather more quickly to the surface. We note that the lithosphere is thinned by a factor two in much shorter time for a thinner lithosphere, indicating some sort of positive feedback process. Because of the high Rayleigh number, the small wavelength and the localized nature of θ and w below the lithosphere, this subscale convective process would indeed be difficult to simulate with the full set of hydrodynamical equations.

Figure 3 shows the time history of the thickness of the lithosphere, defined by the depth of the 1,300 °C isotherm, for various values of T_b , which are all below the solidus³⁰ of dry olivine. The dashed curves represent an initially thin lithosphere with a thickness of 100 km, whereas the solid curves are associated with a thicker lithosphere of 200 km. The wavelength

is maintained at 150 km. For $T_b = 1,550^\circ\text{C}$, and a thin lithosphere, the thinning rate is between 5 and 10 km Myr⁻¹; the rates for a 200-km-thick lithosphere are half as large. One observes that the rate does not decrease with time but rather is constant. This is because the low viscosity channel becomes more pronounced in a thinner lithosphere. Consequently, the local effective Rayleigh number is greatly enhanced, giving rise to convective heat transport that is still able to overwhelm the increased conductive loss across a thinner plate. Under realistic conditions in the Earth, this type of thinning process must stop when the migrating boundary encounters a layer of intrinsically lower density, such as the crust, or a layer of depleted mantle material situated in the top 50 km of the oceanic lithosphere.

The relative uplifts accompanying the lithospheric thinning process are shown in Fig. 4. Elevations of 2 km are attained within 50 Myr for T_b greater than $1,550^\circ\text{C}$ and a thick lithosphere; more rapid rates of uplifting are found for thin lithospheres as a consequence of lower asthenospheric viscosity. We have calculated the elevation changes from thermal variations in the top 200 km of the cell. Uplift rates characteristic of those observed over oceanic swells¹³, $O(0.1 \text{ km Myr}^{-1})$, are attained for values of $T_b > 1,550^\circ\text{C}$.

To use our model for the uplift patterns around hot spots, we must go back to Fig. 1 where the several isotherms depicting the plume are drawn schematically. There one can discern that T_b should be highest at the centre of the plume and should diminish away from it. In Figs 3 and 4 one finds the rate of uplift and thinning varying directly with T_b . Thus one may expect maximum rates of elevation to change and lithospheric thinning to occur in the centre of the upwelling, which is still below the solidus temperature. Quantitative evaluation of how these rates decrease away from the centre requires a truly three-dimensional model. Moreover, the upper surface of our model is stationary, very much unlike the fast-moving Pacific plate. Hence, three-dimensional calculations are deemed necessary to study the interaction of the secondary convection with the shear flow driven by plate motion.

The secondary convection model proposed here can generate the proper time scales for the uplift and thinning, as observed geologically. The viscosity decrease under the lithosphere, which makes rapid thinning by small-scale convection feasible, is a direct consequence of a temperature increase in $\eta(T, p)$. This mechanism becomes more efficient with decreasing viscosity, in contrast to the behaviour exhibited by the large-scale upwelling¹² using $\eta(T)$. However, for higher viscosity values large-scale plumes may compete in the thinning process with small-scale convection.

Any other mechanism able to reduce the viscosity dramatically would produce similar results by means of small-scale flow. This may arise from mantle volatiles containing H₂O or CO₂ being entrained by the plume, as rheological softening may result from small changes in the intrinsic chemical properties of the medium. Hence, the mantle beneath swells may not necessarily be anomalously warm. Finally, we note that this type of small-scale convection may have played a more important role in the global heat-transport during the early history of the Earth—the mantle temperatures were higher then—thus promoting more vigorous rates of lithospheric erosion.

We thank Uli Christensen and Peter M. Olson for helpful criticisms. This research has been supported by NSF grants EAR-8214094 and EAR-8306738; Petroleum Research grant 13550-G2, administered by the American Chemical Society; NATO Research Award (grant no. 27681); and by the Institut National d'Astronomie et de Géophysique (INAG).

Received 3 April; accepted 1 October 1984.

- Yegorin, A. V. & Pavlenkova, N. I. *Phys. Earth planet. Int.* **25**, 12–26 (1981).
- Maguire, P. K. H. & Long, R. E. *Geophys. J. R. astr. Soc.* **44**, 661–675 (1976).
- Spohn, T. & Schubert, G. *J. geophys. Res.* **87**, 4669–4781 (1982).
- Crough, S. T. & Thompson, G. A. *Earth planet. Sci. Lett.* **31**, 397–402 (1976).
- Mareschal, J. C. *Tectonophysics* **94**, 51–66 (1983).
- Morgan, P. *Tectonophysics* **94**, 277–298 (1983).
- Emery, S. H. & Turcotte, D. L. *J. Fluid Mech.* **127**, 507–517 (1983).
- Heestad, R. L. & Crough, S. T. *J. geophys. Res.* **86**, 6107–6114 (1981).
- Detrick, R. S. & Crough, S. T. *J. geophys. Res.* **83**, 1236–1244 (1978).
- Wendlandt, R. F. & Morgan, P. *Nature* **298**, 734–736 (1982).
- Crough, S. T. *A. Rev. Earth planet. Sci.* **11**, 165–193 (1983).
- Sclater, J. G. *Earth planet. Sci. Lett.* **51**, 139–162 (1980).
- Bird, P. *J. geophys. Res.* **84**, 7561–7571 (1979).
- Bridwell, R. J. in *Tectonics and Geophysics of Continental Rifts* (ed. Ramberg, H. B. & Newman, E. R.) 73–80 (Reidel, Dordrecht-Boston, 1977).
- Neugebauer, H. J. *Tectonophysics* **94**, 91–108 (1983).
- Parmentier, E. M., Turcotte, D. L. & Torrance, K. E. *J. geophys. Res.* **80**, 4417–4424 (1975).
- Christensen, U. R. *Ann. Geophysicae* **2**, 311–320 (1984).
- Whitehead, J. A. & Luther, D. S. *J. geophys. Res.* **80**, 705–717 (1975).
- Olson, P. L. & Singer, H. *J. Fluid Mech.* (in the press).
- Richter, F. M. & Parsons, B. E. *J. geophys. Res.* **80**, 2529–2541 (1975).
- Parsons, B. E. & McKenzie, D. P. *J. geophys. Res.* **83**, 4485–4496 (1978).
- Clever, R. M. & Busse, F. H. *J. Fluid Mech.* **65**, 625–645 (1974).
- Gough, D. O., Spiegel, E. A. & Toomre, J. *J. Fluid Mech.* **68**, 695–719 (1975).
- Fleitout, L. & Yuen, D. A. *J. geophys. Res.* **89**, 9227–9244 (1984).
- Kushiro, I., Syono, Y. & Akimoto, S. *J. geophys. Res.* **73**, 6023–6029 (1968).

A fossil hydrothermal worm assemblage from the Tynagh lead–zinc deposit in Ireland

David A. Banks

Department of Applied Geology, University of Strathclyde, Glasgow G1 1XJ, UK

Fossil hydrothermal chimneys were first discovered at Silvermines¹, Ireland. Subsequent discoveries of fossils associated with mounds and chimneys in sulphide deposits within the ophiolite complexes of Cyprus² and Oman³ prove that life forms similar to those colonising active vents on the East Pacific Rise^{4–7} were extant 100 Myr ago around ridge crest hot springs. Here I report the discovery of 350-Myr-old worm-like fossils recovered from pyrite chimneys associated with sedimentary-exhalative mineralization at Tynagh, Ireland. This find suggests that even older examples of life existing at seafloor hot springs may be awaiting discovery.

Modern deep-sea hot springs are features associated with medium to fast ocean-floor spreading centres where the high heat flow, related to shallow magmatic intrusions, drives hydrothermal convection cells within the newly formed oceanic crust. Metals released from the crust during hydrothermal metamorphism on the East Pacific Rise are precipitated around seafloor vents as metalliferous deposits whose spectacular end-members are the black and white smokers at 21°N^4 . Normally the faunal population of the deep oceans is sparse⁸, restricted by the limited supply of nutrient falling as particulate matter from above. However, at hot spring sites in the eastern Pacific^{4–7}, the local biological communities are extraordinarily dense owing to the prodigious increase in the supply of nutrients and include clams, tube worms, crabs, mussels, anemones, barnacles, limpets and many other species. Chemoautotrophic bacteria, which constitute the base of the food chain, use H₂S dissolved in the hydrothermal fluid⁹, as a source of chemical energy. This food chain is probably utilized by the filter-feeding animals (such as serpulid worms) but it is less likely that the giant vestimentiferan worms^{10–12} are nourished by this method. The worms live attached to rocks bathed by water at approximately 20°C and grow to 3 cm in diameter and 3 m in length. They lack a mouth, gut and anus and probably live on dissolved organic matter such as amino and carboxylic acids.

The discovery of these active hydrothermal vents and associated biological communities along the East Pacific Rise^{4–7} has prompted a search for similar faunal occurrences in orebodies held to be the result of sedimentary deposition from hot springs. Structures similar to hydrothermal chimneys and organic remains have been found² in the massive lenticular sulphide

- Crough, S. T. *Geophys. J. R. astr. Soc.* **55**, 451–469 (1978).
- Haxby, W. F. & Turcotte, D. L. *J. geophys. Res.* **83**, 5478–5487 (1978).
- Crough, S. T. *Tectonophysics* **77**, 189–202 (1981).
- Crough, S. T. *Geophys. Res. Lett.* **8**, 877–879 (1981).
- Brott, C. A., Blackwell, D. D. & Ziagos, J. P. *J. geophys. Res.* **86**, 11709–11734 (1981).

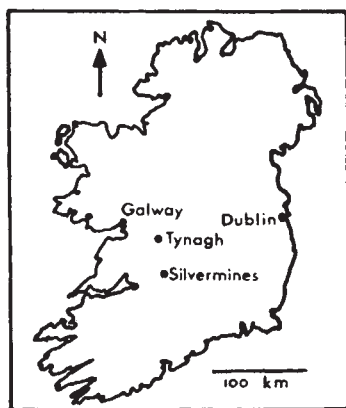


Fig. 1 Map of Ireland with location of Tynagh and Silvermines.

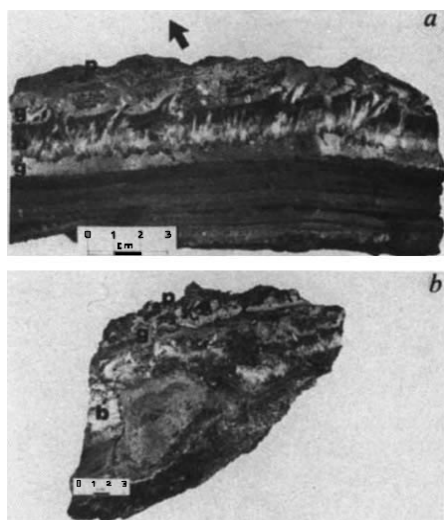


Fig. 2 *a*, Laminated sedimentary sulphides underlying the pyrite mound and chimneys. Arrow indicates the orientation of the central channel of the chimneys. Note the flowering-up structure of the lower galena layer and the imbrication of the upper layer. p, Pyrite; g, galena; b, baryte. *b*, Fracturing and disruption of the layered sulphides beneath the pyrite mound.

deposits associated with the Troodos Ophiolite complex on Cyprus. The 'tiger's eye' structures¹³ found in the Kuroko deposits have been interpreted¹⁴ as being smaller versions of the chimneys at 21° N along the East Pacific Rise. Until now the best examples of the actual fauna supported by hot springs at an ancient oceanic spreading axis were fossils of hydrothermal vent worms, found at Bayda³. Fossil hydrothermal chimneys were first discovered¹ at Silvermines, Ireland, and are much smaller than their modern counterparts, having different morphologies and mineralogies probably owing to the lower flow rates and temperatures of the exhaled solutions¹⁵. Fossil hydrothermal worms and chimneys have now been found at Tynagh.

The Tynagh Pb-Zn-Cu-Ag-BaSO₄-ore deposit¹⁶ is located in the west of the central Irish plain 45 km east south east of Galway (Fig. 1). The deposit lies in a small intracontinental basin adjacent to a fault which was active during deposition of the host rocks. These are micrites and biomicrites of bank facies with impure limestones of off-bank basinal and possibly lagoonal facies. Water depth at the time of mineralization is considered to have been less than 100 m. Mineralization is of Upper Tournaisian-Lower Viséan (Lower Mississippian) age (350 Myr), mainly as replacement and fracture fillings in Waulsortian carbonate mudbanks. Primary mineralization consists of pyrite, galena and sphalerite with some chalcopryrite and tennantite often intimately mixed with baryte. An exhalative banded-iron formation extends for some 500 m to the north of the main

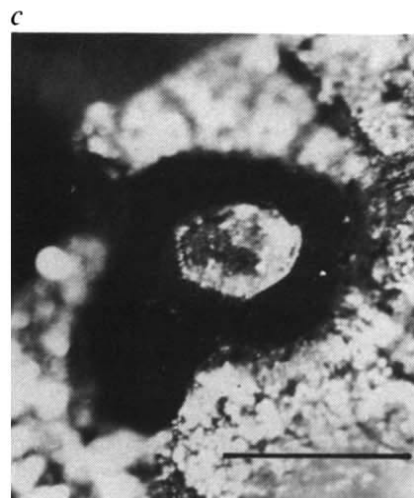
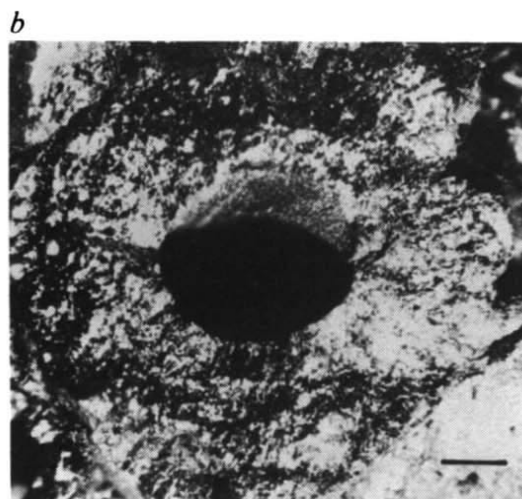
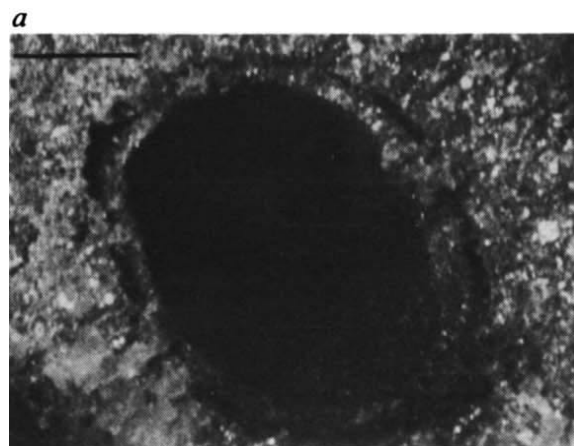


Fig. 3 *a*, Type 1 structure found in the pyrite mound interpreted as a worm-tube mould. *b*, Type 2 structure found in the pyrite mound interpreted as a small chimney produced by exhaling hydrothermal fluids. The walls are much thicker than those in *a*. *c*, Pyrite chimney with a fossil worm *in situ*. Scale bars, 1 mm.

mineralization. A residual ore-body¹⁷ of Tertiary age lies above the more westerly of the two primary mineralized zones. It is postulated that the ore solutions were derived from subsurface circulation of modified seawater initiated during Lower Carboniferous rifting and driven by a high geothermal gradient¹⁸. Boast *et al.*¹⁹ suggest that acid hydrothermal solutions, bearing metals and some primary H₂S, migrated up and along the Tynagh

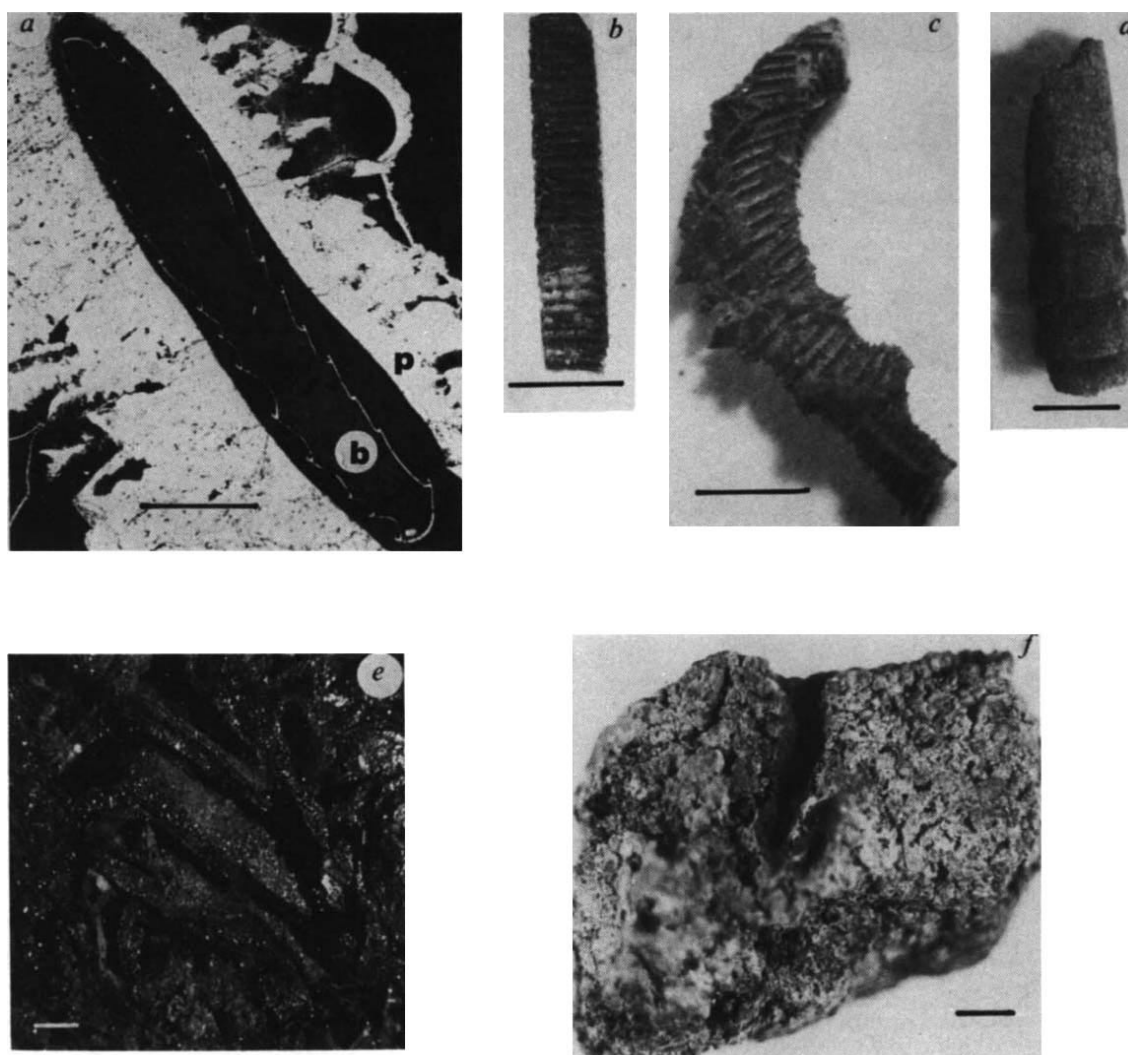


Fig. 4 *a*, Fossil worm enclosed in a pyrite chimney. External structure of the worm is delineated by a fine coating of pyrite. The internal structure of the worm has been completely replaced by baryte. *p*, Pyrite; *b*, baryte. *b*, Fossil worm showing close spaced annuli, 0.1 mm apart. *c*, Curved example of *b*. *d*, Fossil worm with more pronounced annuli, 0.3 mm apart. *e*, Part of a larger segmented animal (possibly a cephalopod) found within massive pyrite. *f*, Part of a worm burrow or a tube mould found in the same massive pyrite as *e*. Scale bars are all 1 mm except *f* which is 10 mm.

Fault and precipitated sulphide minerals, in response to a decrease in temperature and an increase in pH. Additional sulphide was contributed by bacterial reduction of seawater sulphate.

At Tynagh, I envisage the pyrite mound and chimneys to have formed on a fault scarp when hot buoyant iron-bearing solutions rose to the seafloor through small fissures created in response to movement on the Tynagh Fault. Such a view is supported by the chimneys being at an angle of approximately 50° to the sedimentary sulphides as well as by imbrication of the upper galena layer which indicates movement while still semi-lithified; the baryte crystals are also bent in the downslope direction (Fig. 2*a*). The underlying layers of galena, sphalerite, baryte and minor tennantite, typical of the sparse development of sedimentary ore at the deposit, are also fractured and disrupted (Fig. 2*b*).

Surviving chimney structures at Tynagh, composed entirely of pyrite of varying crystal sizes, are up to 15 mm in length, 7 mm in diameter with a central channel of up to 4 mm in diameter. Generally the orifices are circular in plan but a few are elliptical; they may be empty, filled with a later generation of framboidal pyrite or contain the remains of worm-like fossils. Linear ornaments on the inner walls of the tubes are a common feature. Modern hydrothermal chimneys, such as those in the

Pacific at 21° N⁴, grow up to 5 m in height, 30 cm in diameter and are composed primarily of cylindrical shells of Zn, Cu and Fe sulphides together with anhydrite and smaller quantities of silicates²⁰.

Two types of chimney-like structures have been recognized at Tynagh. Type 1 (Fig. 3*a*) has a central aperture approximately 1.5 mm in diameter, walls approximately 0.2 mm thick and no definite crystal structure. These structures are interpreted as worm burrows or, more likely, worm-tube moulds, bearing a strong resemblance to worm-tube moulds recovered from Juan de Fuca²¹, Troodos² and Bayda³. Type 2 (Fig. 3*b*) structures have a central aperture approximately 2 mm in diameter and walls 2 mm thick and a radiating crystal structure. These are interpreted as small chimneys produced by the exhaling hydrothermal fluids, comparable to the type 2*a* chimneys found at Silvermines²², whose abiogenic origin¹ is confirmed by their close-packing, parallelism, absence of cross-cutting structures and the lack of fossil-organic remains.

The remains of numerous vermiform fossils have been discovered within the hydrothermal chimneys on a pyrite mound at Tynagh. Specimens have been found in both of the above-mentioned types of chimney-like structures (Fig. 3*c*). These fossils have a thin outer coating of fine pyrite enclosing baryte

which has completely replaced the internal structure of the organism (Fig. 4a). Maximum dimensions of the recovered fossils are 10 mm in length and a diameter of 0.8 mm. The fossils are usually straight but occasionally, curved examples have been recovered. Two different types of worm have been found: one has closely spaced annuli 0.1 mm apart (Fig. 4b, c), whereas the other has annuli which are more pronounced and 0.3 mm apart (Fig. 4d). These could represent different growth stages of the same kind of worm but this is unlikely as the overall dimensions of both are similar. Two further species were found in massive pyrite which may have formed contemporaneously with the pyrite chimneys. The first is the remains of a larger segmented animal (Fig. 4e), possibly a cephalopod; the other is either a worm burrow or, more likely, part of a worm-tube mould (Fig. 4f). The tube moulds have diameters up to 7 mm with walls 0.5 mm thick. Their means of preservation could have been similar to that of the present-day tube moulds²¹ found on the East Pacific Rise, where exhaled sulphides rain down and cover animals on the seafloor. It may be that particular hydrothermal vents were colonized by worms, during periods of quiescence, and subsequent reactivation of the hot fluid caused their demise. Closure of the vents effectively entombed the worms resulting in their well-preserved state.

Why did the hydrothermal chimneys support metazoan life at Tynagh whilst the vents at Silvermines of exactly the same age, and only 50 km to the south, are devoid of such fossils? Most of the zinc and lead at Silvermines was mined from a sedimentary exhalative pyrite horizon. At Tynagh the exhalative iron-rich sediment is in oxide facies and there is little sedimentary sulphide ore; most of the lead and zinc was mined from shallow replacements in limestone. To explain these differences Russell²³ has argued that the sedimentary sulphides at Silvermines were protected from oxidation by being laid down in a brine pool, whereas at Tynagh hydrothermal solutions exhaled into an oxidizing Carboniferous sea. The most likely origin for the dense brine is highly saline seawater²⁴, produced by evaporation on the margin of a fault-controlled basin²⁵ which occupied the Silvermines area during the Lower Carboniferous. Fluid inclusion results²⁴ are supportive of the presence of such a brine during mineralization. Boyce *et al.*²² have argued that conditions at Silvermines were too saline to support life, whereas at Tynagh brine pools did not exist so the less saline and more oxidizing conditions were favourable to life. Nevertheless, the type 2 mounds and vents at Silvermines, on the southern periphery of the mineralized basin, occupy a high structural position above the halocline and I suggest a fauna may yet be discovered there.

The fossils of hydrothermal vent worms from the Bayda sulphide deposit³ are less well preserved than the specimens from Tynagh. Features such as size, closely-spaced annulations on the tube casts, and the presence of worm-tube moulds suggest that the two faunal populations were broadly similar. The tube moulds from Tynagh, Bayda³ and Troodos² are similar to present-day tube moulds recovered from hot springs on the East Pacific Rise, especially those from Juan de Fuca²¹. So far, clams, crabs and giant tube worms, evident at modern hot springs, have not been found in the ancient deposits. The Tynagh worms themselves are probably best compared with the smaller polychaete worms which form dense encrustations around the hot springs along the East Pacific Rise. At present it is impossible to tell if the Tynagh worms form part of an ancestral line to modern worms, or if repeated readaptation to the hot-spring environment has taken place. Certainly it would not be difficult to find a comparable modern species purely on the basis of external form, the polychaetes alone have over 5,300 different species. What is significant, however, is that similar fauna have been found in hydrothermal environments of various ages and we may therefore expect to find fauna at even older vent sites.

I thank the staff and postgraduates of the Applied Geology Department, especially John Gilleece, Mike Russell, George Bowes and Adrian Boyce; John Clifford (Ennex International) for his help and permission to work on the Tynagh deposit; and

Strathclyde University for financial support by the award of a John Anderson Postgraduate Studentship.

Received 5 July; accepted 30 October 1984.

1. Larter, R. C. L., Boyce, A. J. & Russell, M. J. *Mineralium Deposita* **16**, 309–318 (1981).
2. Oudin, E. & Constantinou, G. *Nature* **308**, 349–352 (1984).
3. Haymon, R. M., Koski, R. A. & Sinclair, C. *Science* **223**, 1407–1409 (1984).
4. Speiss, F. N. *et al. Science* **207**, 1421–1433 (1980).
5. Corliss, J. B. *et al. Science* **203**, 1073–1082 (1979).
6. Corliss, J. B. & Ballard, R. D. *Natn. Geogr.* **152**, 440–453 (1977).
7. Ballard, R. D. & Grassle, J. F. *Natn. Geogr.* **156**, 680–705 (1979).
8. Banse, K. *Prog. Oceanogr.* **2**, 53–125 (1964).
9. Edmond, J. M. *et al. Earth planet Sci. Lett.* **46**, 1–18 (1979).
10. Jones, M. L. *Proc. biol. Soc. Wash.* **93**(4), 1295–1313 (1980).
11. Jones, M. L. *Science* **213**, 333–336 (1981).
12. Felbeck, H. *Science* **213**, 336–338 (1981).
13. Narita, S. S. *et al. Min. Geol.* **27**(2), 119–132 (1977).
14. Scott, S. D. *Geosci. Canada* **8**(3), 103–104 (1981).
15. Russell, M. J., Boyce, A. J., Larter, R. C. L. & Samson, I. M. in *Mineral Exploration in Ireland*, 171–172 (Irish Association for Economic Geology, 1981).
16. Derry, C. J., Clark, G. R. & Gillatt, N. *Econ. Geol.* **60**, 1218–1237 (1965).
17. Morrissey, C. J. & Whitehead, D. *Proc. 9th Commonwealth Min. Metall. Congr.* Vol. 2, 131–145 (London Institution of Mining and Metallurgy, 1970).
18. Russell, M. J., Solomon, M. & Walshe, J. L. *Mineralium Deposita* **16**, 113–127 (1981).
19. Boast, A. M., Coleman, M. L. & Halls, C. *Econ. Geol.* **76**, 27–55 (1981).
20. Haymon, R. M. & Kastner, M. *Earth planet Sci. Lett.* **53**, 363–381 (1981).
21. Koski, R. A., Clague, D. A. & Oudin, E. *Bull. geol. Soc. Am.* **95**, 930–945 (1984).
22. Boyce, A. J., Coleman, M. L. & Russell, M. J. *Nature* **306**, 545–550 (1983).
23. Russell, M. J. in *Short Course in Sediment Hosted Stratiform Lead–Zinc Deposits* Vol. 8, 252–282 (Mineralogical Association of Canada, 1983).
24. Samson, I. M. & Russell, M. J. *Trans. Instn Min. Metall.* **B92**, 67–71 (1983).
25. Boyce, A. J., Anderton, R. & Russell, M. J. *Trans. Instn Min. Metall.* **B92**, 55–66 (1983).

Parental investment and sex differences in juvenile mortality in birds and mammals

T. H. Clutton-Brock, S. D. Albon & F. E. Guinness

Large Animal Research Group, Department of Zoology, University of Cambridge, Cambridge CB3 0DT, UK

The common finding that juvenile male mammals show higher mortality than females^{1,2} is usually attributed either to the expression of deleterious recessive alleles on the X chromosome³ or to adaptive manipulation of the postnatal sex ratio by mothers who are unable to rear successful sons^{4,5}. As a general explanation of differential juvenile mortality, the first of these two theories is unsatisfactory because increased male mortality is known to occur in several bird species where females are the heterogametic sex^{6,7}. The second hypothesis was first proposed by Trivers and Willard⁴ and suggests that differential mortality should occur early in the period of parental investment⁸. We now show that, although several predictions of the latter explanation are fulfilled, the distribution of differential mortality in red deer and other mammals suggests that higher mortality rates among male juveniles are a consequence of a greater susceptibility of males to food shortage associated with their faster growth rates and increased nutritional requirements.

Patterns of mortality in red deer (*Cervus elaphus*) on the island of Rhum⁹ appear to support Trivers and Willard's suggestion that sex differences in juvenile mortality occur because inferior mothers prematurely terminate investment in their sons. As we have shown previously¹⁰, the sons of subordinate hinds rarely breed successfully, while their daughters are almost as successful as those of dominant mothers. According to Trivers and Willard's hypothesis, subordinate mothers might be expected to reject their sons, and sex differences in juvenile mortality should be confined to their offspring. As predicted, we found that the sons of subordinate mothers are more likely to die than are their daughters, but that there is no sex difference in mortality among the offspring of dominant hinds (Table 1).

The theory also suggests that mortality among juvenile males relative to that among juvenile females should rise as resource availability declines. Since 1971, the number of hinds ≥ 1 year old using our study area has increased from 57 to 166, feeding competition has increased, and most measures of reproductive performance have declined⁹. As expected from the theory, the mortality rate among juvenile males has risen relative to that

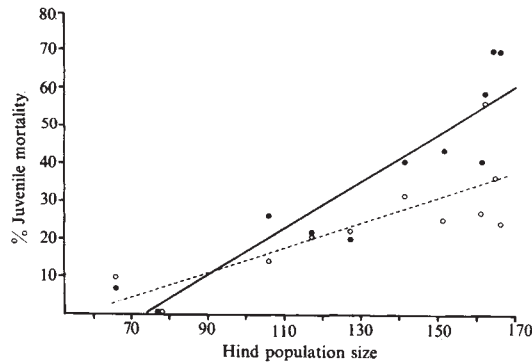


Fig. 1 Mortality during the first 2 yr of life among males (●) and females (○) born between 1971 and 1982 in the red deer population of the North Block of Rhum (regression equations: males, $y = 0.622x - 45.45$, $r = 0.913$; $F_{1,9} = 43.8$, $P < 0.001$; females, $y = 0.328x - 18.487$, $r = 0.788$; $F_{1,9} = 12.69$, $P < 0.01$. Difference in slopes, $F_{1,18} = 4.96$, $P < 0.05$).

among juvenile females (Fig. 1). Similar results have been reported in wood rats⁵.

Finally, as maternal investment and early growth are likely to have the greatest influence on male breeding success in polygynous species showing pronounced sexual dimorphism in body size^{4,9}, sex differences in juvenile mortality should be most frequently found in these species. Although reliable evidence of differential juvenile mortality in mammals is rare, partly because of sex differences in juvenile visibility¹¹ and dispersal¹², the available evidence suggests that it is usually slight or absent in monomorphic species and is most pronounced in polygynous species where males are substantially larger than females. Figure 2a shows, for nine ungulate species, the difference between estimates of the sex ratios (% males) at or soon after birth and towards the end of the first year of life. A similar relationship between differential juvenile mortality and adult sexual dimorphism may exist among birds during the interval between hatching and fledging (Fig. 2b). These figures must be interpreted cautiously since samples were small and, in some cases, estimates of the sex ratio at different stages were drawn from different studies. In addition, there may be some exceptions: in great tits, one study found that the sex ratio of nestlings was male biased by 15 days post hatching¹³.

Although these results seem to support Trivers and Willard's adaptive explanation of differential mortality, all three trends may be explained in another way. In sexually dimorphic birds and mammals, males grow faster and for longer than females and require more food¹⁴⁻¹⁵. For example, nestling male red-winged blackbirds eat 20–30% more than females of the same weight¹⁶; among feral goats, male kids suck more than twice as frequently as females¹⁷; and in domestic sheep, young rams require 15% more digestible nutrients than ewes of equal weight and age¹⁸. Sex differences in growth and metabolism are associated with differences in growth priorities: among birds, the smaller sex commonly fledges more rapidly¹⁹ and juvenile male mammals lay down less body fat than females⁹. Across species, sex differences in growth are closely related to the degree of adult sexual dimorphism in body size, and differences in metabolic requirements and body fat probably follow the same trend. For example, the ratio of male to female fatness among juveniles of different deer species is negatively correlated with the ratio of adult male to female body weight (Table 2: $r_s = -0.717$, $n = 9$, $P < 0.05$). In the only species in which females are larger than males (Chinese water deer; S.D.A., unpublished data), the direction of the sex difference in fatness is reversed. And in a wide variety of dimorphic mammals there is evidence that food shortage affects the growth rates of males more than females⁹.

Thus, a reduction in the viability of males when food is short, resulting from their greater nutritional requirements, could account for (1) the relationships between differential mortality and sexual dimorphism (Fig. 2), (2) the tendency for sex differences in mortality to increase as population density increases

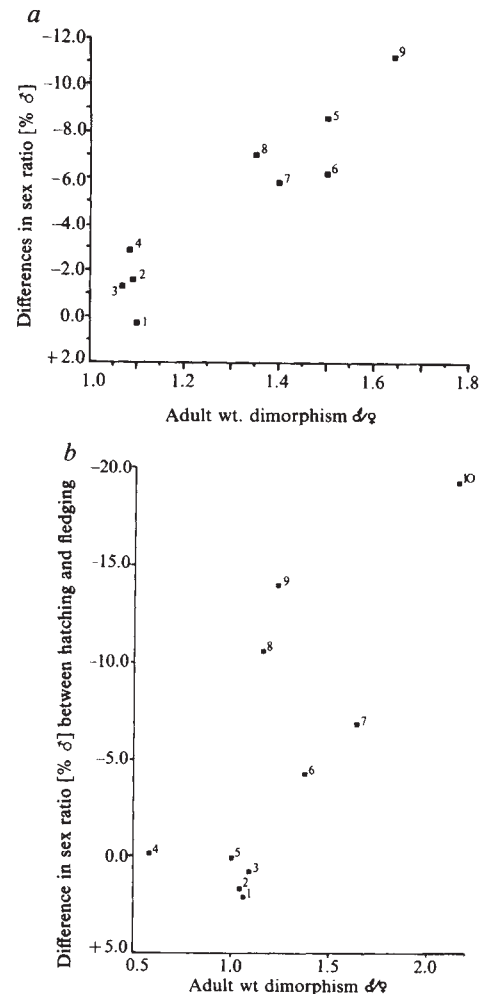


Fig. 2 a, Differences in the sex ratio (% males) between birth and the end of the first year of life in free-ranging populations of ungulates, showing different degrees of adult weight dimorphism ($\bar{x}$ male weight/ $\bar{x}$ female weight). 1, Zebra²⁸; 2, roe deer¹¹; 3, sable²⁹; 4, feral horse (P. Duncan, personal communication); 5, black-tailed deer³⁰; 6, red deer⁹; 7, wapiti^{31,32}; 8, Soay sheep³³; 9, reindeer^{34,35}. b, Differences in the sex ratio (% males) between hatching and fledging in bird species showing different degrees of adult weight dimorphism ($\bar{x}$ male weight/ $\bar{x}$ female weight). 1, Eastern bluebird³⁶; 2, starling³⁷; 3, snowgoose^{38,39}; 4, European sparrowhawk²²; 5, red-cockaded woodpecker (P. A. Gowaty and M. R. Lennartz, in preparation); 6, red-winged blackbird^{16,40}; 7, yellow-headed blackbird^{19,41}; 8, rook⁴²; 9, common grackle⁷; 10, capercaillie⁴³.

and food availability declines (Fig. 1) and (3) the restriction of differential mortality to the offspring of inferior mothers (Table 1). The evolutionary explanation of differential mortality may be that sexual selection favouring large adult size in males has increased male growth rates to a point at which costs to juvenile survival during periods of food shortage balance the benefits of large size in adulthood.

The predictions of this theory differ from those of Trivers and Willard's hypothesis in at least three ways. Sex differences in mortality should not be confined to the early stages of parental investment and would be expected throughout the period of differential growth; they should not necessarily be associated with a reduction in parental investment in juveniles of one sex; and they should persist when juveniles are reared on inadequate diets in the absence of their parents.

The available evidence is consistent with an explanation of differential mortality based on sex differences in growth and does not support Trivers and Willard's explanation. Among many ungulates, differential mortality is most pronounced during the first and second winters, after juveniles have been weaned

Table 1 Percentage mortality of male and female red deer calves (0–12 months) and yearlings (12–24 months) in relation to the dominance rank of their mothers

Mother's rank	Calves		P	Yearlings		P
	Males	Females		Males	Females	
Dominant	36 (185)	36 (125)	NS	8 (115)	11 (80)	NS
Subordinate	51 (102)	30 (109)	<0.001	30 (50)	12 (76)	<0.05

Sample sizes are given in parentheses. Dominant hinds were those above median rank, subordinates those below it¹⁰. NS, Not significant.

Table 2 Sex differences in $\bar{x}$ adult body weight (kg) and in juvenile fatness (kidney fat index) in winter in different deer species

	Adult body weight (kg)		Ratio (male/female)	Ratio of kidney fat indices for male and female juveniles (male/female)
	Male	Female		
Chinese water deer <i>Hydropotes inermis</i>	11.3	13.3	0.85	2.64
Reeve's muntjac <i>Muntiacus reevesi</i>	13.5	11.8	1.16	0.98
Roe deer <i>Capreolus capreolus</i>	26	24	1.09	0.79
Fallow deer <i>Dama dama</i>	64	46	1.39	0.87
Mule deer <i>Odocoileus hemionus</i>	90	64	1.41	0.67
White-tailed deer <i>O. virginianus</i>	96	62	1.55	0.45
Red deer <i>Cervus elaphus</i>	122	81	1.52	0.72
Wapiti <i>Cervus canadensis</i>	340	260	1.31	0.55
Reindeer <i>Rangifer tarandus</i>	140	93	1.56	0.73

Data extracted from the literature (N. Chapman, unpublished). The kidney fat index is the weight of kidneys and perinephric fat divided by kidney weight and is a reliable measure of condition²⁷.

(see refs in Fig. 2a): for example, in the red deer of Rhum, the % ratio of male to female mortality between birth and 6 months (when most offspring are weaned) is 1.09, compared with 1.33 during the second 6 months of life and 1.58 during the second year of life (see Table 1). Although some studies have suggested that sex differences in mortality are caused by maternal rejection of feeding attempts by one sex^{5,7}, no firm evidence for this has yet been published. And experiments involving the maintenance of juvenile rats and pigs on inadequate diets show that higher rates of mortality among males occur in the absence of their parents²⁰.

There are, however, several anomalies which suggest that other factors are involved besides the faster growth rates of males and their greater nutritional requirements. These include a small but consistent tendency for males to show higher hatching or neonatal mortality in some species showing little size dimorphism (including thoroughbred horses and man^{1,21}): the absence of evidence that the direction of differential mortality is reversed in species where females are larger than males (although adequate data are so far available only for the European sparrowhawk²²); and the tendency for male fetuses to be less viable than females during the early stages of gestation^{1,23–25}.

The existence of increased mortality among male juveniles relative to females indicates that the common assumption that juvenile sex ratios approximate to parity may often be incorrect and may help to explain why adult sex ratios are often biased

towards females in mammal populations maintained close to carrying capacity^{9,26}.

We thank the Nature Conservancy Council for permission to work on Rhum; Glenn Iason, Callan Duck and Martin Major for their collaboration in the research on red deer; N. Chapman and P. Gowaty for access to unpublished data; and Nick Davies and David Gibbons for their comments. The research is funded by grants from the NERC, SERC and the Royal Society.

Received 20 June; accepted 9 October 1984.

1. Parkes, A. S. *Biol. Rev.* **2**, 1–51 (1926).
2. Ralls, K., Brownwell, R. L. & Ballou, J. *Rep. int. Whal. Comm. Spec. Issue*, **2**, 223–243 (1980).
3. Myers, J. H. *Am. Nat.* **112**, 381–388 (1978).
4. Trivers, R. L. & Willard, D. E. *Science* **179**, 90–92 (1973).
5. McClure, P. A. *Science* **211**, 1058–1060 (1981).
6. MacArthur, J. W. & Baillie, W. H. T. *Q. Rev. Biol.* **8**, 313–325 (1932).
7. Howe, H. *Science* **198**, 744–746 (1977).
8. Maynard Smith, J. *Behav. Ecol. Sociobiol.* **7**, 247–251 (1980).
9. Clutton-Brock, T. H., Guinness, F. E. & Albon, S. D. *Red Deer: Behavior and Ecology of Two Sexes* (University of Chicago Press, 1982).
10. Clutton-Brock, T. H., Albon, S. D. & Guinness, F. E. *Nature* **308**, 358–360 (1984).
11. Anderson, J. *Dan. Rev. Game Biol.* **2**, 127–155 (1953).
12. Greenwood, P. J. *Anim. Behav.* **28**, 1140–1162 (1980).
13. Dhondt, A. *Bird Study* **17**, 282–286 (1970).
14. Glucksman, A. *Biol. Rev.* **49**, 423–475 (1974).
15. Case, T. J. *Q. Rev. Biol.* **53**, 243–282 (1978).
16. Fiala, K. L. In *Natural Selection and Social Behavior* (eds Alexander, R. D. & Tinkle, D. W.) (Chiron, New York, 1981).
17. Pickering, S. P. thesis, Univ. Durham (1983).
18. Morrison, J. *Foods and Feeding* 4th edn (Rheingold, New York, 1948).
19. Richter, W. *Am. Nat.* **121**, 158–171 (1983).
20. Widdowson, E. M. *Proc. Nutr. Soc.* **35**, 175–176 (1976).
21. Platt, H. *A Survey of Perinatal Mortality in the Thoroughbred* (Animal Health Trust, Newmarket, 1978).
22. Newton, I. *Population Ecology of Raptors* (Poyser, Berkhamstead, 1979).
23. Chapman, A. B., Casida, L. E. & Cote, A. *Proc. Am. Soc. Anim. Prod.* **1938**, 303–304 (1938).
24. Parkes, A. S. *J. agric. Sci.* **15**, 15–30 (1925).
25. McMillen, M. M. *Science* **204**, 89–91 (1979).
26. Cowan, I. *McT. Trans. N. Am. Wildl. Conf.* **15**, 581–588 (1950).
27. Riney, T. *N.Z. Sci. Tech. B36*(5), 429–463 (1955).
28. Smuts, G. L. S. *Afr. J. Wildl. Res.* **6**, 99–112 (1976).
29. Grobler, J. H. *Arnoldia* **7**, 1–36 (1974).
30. Taber, R. D. & Dasmann, R. F. *Ecology* **38**(2), 233–246 (1954); *J. Wildl. Mgmt* **18**, 309–315 (1954).
31. Flook, D. R. *Canadian Wildlife Service Report Ser. 11* (Department of Indian Affairs and Northern Development, Ottawa, 1970).
32. Houston, D. B. *The Northern Yellowstone Elk* (Macmillan, New York, 1982).
33. Grubb, P. In *Island Survivors: The Ecology of the Soay Sheep of St Kilda* (eds Jewell, P. A., Milner, C. & Boyd, J. M.) 242–273 (Athlone, London, 1974).
34. Bergerud, A. T. *Wildl. Monogr.* **25**, 1–55 (1971).
35. Leader Williams, N. *J. Wildl. Mgmt* **44**, 640–657 (1980).
36. Pinkowski, B. C. *Condor* **79**, 289–302 (1977).
37. Kessel, B. A. *Midl. Nat.* **58**, 257–331 (1957).
38. Ankney, C. D. *Auk* **99**, 662–666 (1982).
39. Harmsen, L. C. & Cooke, F. *Am. Nat.* **122**, 1–8 (1983).
40. Holcomb, L. C. & Twiest, G. *Wilson Bull.* **82**, 294–303 (1970).
41. Patterson, C. B. & Emlen, J. M. *Am. Nat.* **115**, 743–747 (1980).
42. Roskaft, E. thesis, Univ. Trondheim (1983).
43. Wegge, P. *Ornis. Scand.* **11**, 106–109 (1980).

Primate cortical area V4 important for colour constancy but not wavelength discrimination

H. M. Wild, S. R. Butler*, D. Carden & J. J. Kulikowski

Visual Sciences Laboratory, Ophthalmic Optics Department, University of Manchester Institute of Science and Technology, PO Box 88, Manchester M60 1QD, UK

*Anatomy Department, The Medical School, University of Birmingham, PO Box 363, Birmingham B15 2TT, UK

One of the remarkable features of vision is that the perceived colour of reflecting surfaces which form part of complex, multi-coloured scenes, remains stable despite large variations in the spectral composition of the illuminating light^{1–5}. The phenomenon of colour constancy has long been known but only recently have electrophysiological studies revealed a possible mechanism for it in an area of visual cortex, the V4 complex in the rhesus monkey. Lesions in the V4 complex of experimental animals^{6,7} have produced little if any impairment of wavelength discrimination, but

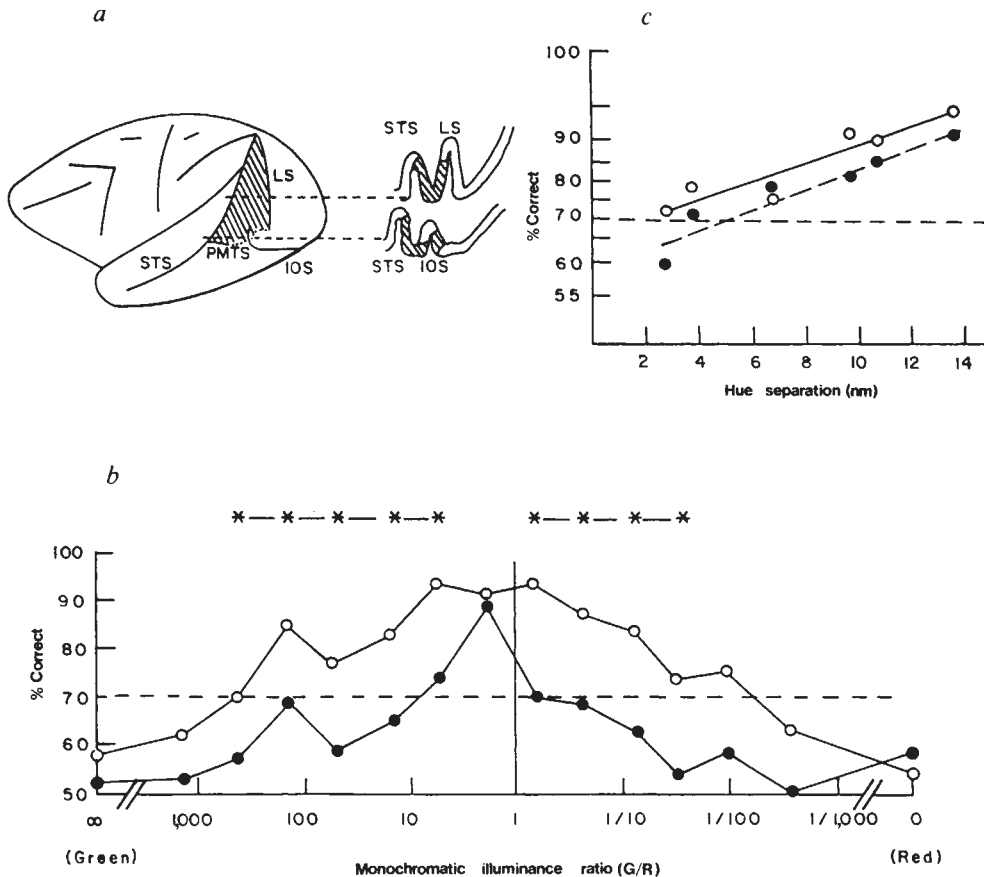


Fig. 1 *a*, Intended extent of V4 ablations in the two macaque monkeys used. LS, lunate sulcus; IOS, inferior occipital sulcus; STS, superior temporal sulcus; PMTS, posterior middle temporal sulcus. *b*, Performance of animals with lesions (●) and controls (○) on the constancy task. Scores are the percentage of correct responses when the animals were required to find the checkerboard plate with a green patch, for 15 different ratios of green/red (G/R) illumination. The horizontal dashed line indicates the level at which performance departs from chance with a confidence limit of $P < 0.01$. * Indicates a significant difference ($P < 0.01$). *c*, Mean scores on the wavelength discrimination task, showing the percentage of correct responses when the animals were required to choose the greener of two plates of very similar hue. Straight lines through the points for each group were fitted by probit analysis. Confidence level as for *b*.

colour constancy has not been tested. In the present study we used a broader range of behavioural tasks, including one of colour constancy, to test the prediction that ablation of the V4 complex leads to selective deficits in colour perception. The results support the notion that the V4 complex is necessary for colour constancy but not for the discrimination of hue.

Cells have been found in the V4 complex which respond to the natural colour of a stimulus irrespective of its local wavelength composition. When a stimulus is illuminated by light comprising the trichromatic components red, green and blue, cells continue to respond as the ratio of the components of the incident illumination is varied over a wide range. The range corresponds to that over which a normal human observer continues to perceive the natural colour⁸, that is, within the range of colour constancy. Other areas of the visual cortex do not show such properties: the V4 complex is flanked by areas V3A and V5 whose cells are not sensitive to chromatic stimuli^{9,10}.

The V4 complex receives its main input from V2 and a smaller contribution from foveal V1, which so far have been found to have only wavelength (not colour) specificity¹⁰⁻¹⁵.

The rhesus visual system resembles that of man, having similar trichromatic colour vision, visual acuity and analogous pathways leading to the primary (V1) visual cortex¹⁶. Defects in colour perception following posterior cerebral lesions in man may take the form of achromatopsia, in which the patient has difficulty distinguishing between similar colours¹⁷⁻¹⁹, or colour agnosia, when the patient is unable to associate objects with their normal colours^{20,21}. To examine the possibility that damage to the V4 complex might be responsible for analogous disorders in monkeys, tests of fine wavelength discrimination, conditional reaction (object-colour association) and short-term memory for colour were performed.

A preliminary series of tests involved training four immature rhesus monkeys on a series of discrimination learning tasks in a modified Wisconsin test apparatus. The stimuli were presented on rectangular plates covering food wells and the well beneath

the 'correct' stimulus was baited with a peanut on each trial. The animals were trained on the following tasks until they reached a criterion of 27 correct trials out of 30 consecutive trials: (1) brightness discrimination; (2) two-dimensional and three-dimensional shape discrimination; (3) discrimination of coloured plaques, and of plaques bearing coloured 'Mondrian' patterns, with brightness and saturation varied randomly to ensure that hue provided the salient cue; (4) a series of conditional reactions in which a choice between two shapes was cued by a third stimulus: brightness, hue or pattern; (5) delayed non-matching to sample to test short-term memory²² for hue and pattern.

Two animals were then subjected to bilateral ablations of the V4 complex by subpial aspiration. Figure 1*a* shows the intended extent of the removal. Histological verification awaits completion of further training. After allowing 3 weeks for recovery, the operated animals and the unoperated controls were tested for immediate retention of each task in three sessions of 30 trials each. They were then re-trained on each task in turn.

All animals relearned the tasks, in most cases in fewer trials than they required for preoperative learning. One of the operated animals needed more trials than the controls to relearn each task; the other animal in this group could not be distinguished from the controls. Neither of the operated animals was selectively impaired on tasks involving coloured cues.

The animals were therefore trained postoperatively on two further tests for colour constancy and then fine wavelength discrimination. Eight plates were prepared bearing a checkerboard pattern of coloured patches varying randomly in brightness and saturation. Three plates carried one green and 11 orange checks, and five contained 12 orange checks. Two plates were presented on each trial and the animals were trained to choose the one which included a green check. The animals were first trained with the plates illuminated by tungsten light, then yellow light was used, comprising a 1:1 mixture of red and green narrow-band monochromatic lights produced by filtering tung-

sten sources with 630-nm and 515-nm interference filters. The green patch was easily detectable in the yellow illumination, but when either source was used on its own, both plates appeared to a human observer as a set of monochromatic patches of varying degrees of lightness; however, on addition of a very small proportion of the second colour to the illumination, the green patch became readily discernible. This was evidently dependent on colour constancy as both green and orange checks, viewed in isolation through a reduction screen, assumed the hue of the predominant light source²³.

Having mastered the task in the yellow light, the animals were tested in 50 trials at each of 15 different red/green illumination ratios, presented in random order. Figure 1b shows the performance of each group of animals under the 15 conditions of illumination. Analysis of variance reveals that the performance of the two groups was not significantly different in monochromatic light (red or green), where all animals scored at chance levels, nor when the incident illumination contained approximately equal proportions of red and green, that is, in yellow light. However, at almost all other levels where performance was greater than chance, the two groups differed significantly ($P < 0.01$). The performance of the animals with V4 lesions deteriorated much more rapidly than that of controls as the proportion of red and green illumination became unequal, on changing the incident illumination from yellow through to red or green. In other words, the 'intact' animals were able to distinguish the green (G) patch even when the incident illumination had changed from yellow to either red (R) (G/R ratio = 1:100) or green (G/R ratio = 100:1) and this ability was impaired in the operated animals. The behaviour of these animals is therefore consistent with Zeki's inference that the cells of area V4 have a special role in the stability of perceived colour^{8,11}. The deficit shown by the operated animals cannot be accounted for in terms of a decrease in visual acuity as the spatial frequency of the patterns used in the conditional reaction (on which the operated animals were not selectively impaired) was higher than that of the checkerboard used in the constancy task.

The results leave open the possibility that the impairment produced by ablation of the V4 complex is not specific to colour constancy and that a deficit appeared because of a difficulty in discriminating small differences in wavelength²⁴; to control for this and for the possibility that the groups were inadequately matched on the basis of preoperative tests, the animals were trained on a task which involved difficult judgements of chromatic difference but under tungsten illumination. Four plates were prepared in single shades of dark yellow, differing in Munsell hue²⁵ by small increments (equivalent Munsell values: (1) 7.5YR, 2/3, (2) 10YR, 2.5/2, (3) 2.5Y, 2.8/2, (4) 5Y, 2.5/2; dominant wavelength 580 ± 7 nm). Two plates were presented on each trial and the animals were trained to choose the 'greener' one. Figure 1c shows that performance ranged from close to chance to more than 90% correct for all animals over the range of wavelength differences. Analysis of variance reveals no significant difference between the groups overall or at any level of task difficulty, suggesting that the impairment of the operated group on the constancy task was due not to difficulty in discriminating small differences in wavelength but to a failure to judge the colour of an area under incident illumination of different spectral compositions.

As V4 removal does not impair the discrimination of simple patterns and objects, nor the use of such cues in conditional reaction and short-term memory, it appears that the underlying perceptual processes do not depend on serial analysis staging through this region. If we have been successful in removing a single functional area, the performance of these tasks is independent of colour constancy. In such a case, information concerning wavelength composition is extracted before the processes of colour constancy and the detection of chromatic contours can operate. Moreover, the mechanism of colour constancy appears to operate in parallel with that of chromatic form detection.

We thank Messrs R. Morrissey and J. Simpson for technical assistance. S.R.B. thanks the Rowbotham Fund for support. H.M.W. is a former Wellcome Trust scholar.

Received 2 July; accepted 22 October 1984.

1. Young, T. *Lectures in Natural Philosophy* Vol. 2 (J. Johnson, London, 1807).
2. Hering, E. *Outlines of a Theory of the Light Sense* (transl.) (Harvard Univ., Cambridge, Massachusetts, 1964).
3. Land, E. *Proc. natn. Acad. Sci. U.S.A.* **45**, 115-129 (1959).
4. Jameson, D. & Hurvich, L. M. *J. opt. Soc. Am.* **49**, 890-898 (1959).
5. Land, E. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5163-5169 (1983).
6. Dean, P. *Expl Brain Res.* **35**, 69-83 (1979).
7. Fries, W. & Zeki, S. *Proc. EBS Workshop. Animal and Human Psychophysics*, 226 (1980).
8. Zeki, S. *Nature* **284**, 412-418 (1980).
9. Zeki, S. *J. Physiol., Lond.* **277**, 273-290 (1978).
10. Zeki, S. *Neuroscience* **9**, 741-764 (1983).
11. Zeki, S. *Brain Res.* **34**, 19-35 (1971).
12. Zeki, S. *J. Physiol., Lond.* **277**, 227-244 (1978).
13. Livingstone, M. S. & Hubel, D. H. *J. Neurosci.* **4**, 309-356 (1984).
14. Gouras, P. J. *J. Physiol., Lond.* **238**, 583-602 (1974).
15. Michael, C. J. *Neurophysiol.* **41**, 1223-1248 (1978).
16. Hubel, D. H. & Wiesel, T. N. *Proc. R. Soc. B* **198**, 1-59 (1974).
17. Meadows, J. B. *Brain* **97**, 615-632 (1974).
18. Damasio, A., Yamada, T., Damasio, H., Corbett, J. & McKee, J. *J. Neurol.* **30**, 1064-1071 (1980).
19. Pearlman, A. L., Birch, J. & Meadows, J. C. *Ann. Neurol.* **5**, 253-261 (1979).
20. Scotti, B. & Spinner, H. *J. Neurol. Neurosurg. Psychiat.* **33**, 22-28 (1970).
21. Varney, N. R. *J. Neurol. Neurosurg. Psychiat.* **45**, 240-252 (1982).
22. Owen, M. J. & Butler, S. R. *Behav. Brain Res.* **3**, 115-123 (1981).
23. Land, E. *Proc. R. Instn Gt Br.* **47**, 23-57 (1974).
24. De Valois, R. L., Morgan, H. C., Polson, M. C., Mead, W. R. & Hull, E. M. *Vision Res.* **14**, 53-67 (1974).
25. *Munsell Book of Colour* (Munsell Colour Co., Baltimore, 1976).

Fetal occipital cortical neurones transplanted to the rostral cortex can extend and maintain a pyramidal tract axon

Brent B. Stanfield & Dennis D. M. O'Leary

The Salk Institute for Biological Studies and The Clayton Foundation for Research—California Division, PO Box 85800, San Diego, California 92138, USA

In adult rats, cortical neurones that send axons through the pyramidal tract are confined to layer V, over the rostral two-thirds of the cerebral hemisphere¹⁻³. However, during the first postnatal week, many neurones in layer V in the occipital cortex (including the visual cortex) also extend axon collaterals through the pyramidal tract and into the spinal cord^{2,4-6}. These occipital corticospinal collaterals are completely eliminated over the subsequent 2 weeks⁶, although their cells of origin do not die². We now report that when portions of the occipital cortex from fetal rats are transplanted to more rostral cortical regions of newborn rats, some of the transplanted neurones not only extend axons through the pyramidal tract, but also maintain these axons beyond the stage at which they are normally eliminated. These results suggest that normally-eliminated cortical axons can be 'rescued' and, in the case of pyramidal tract neurones, the position of the neurones within the tangential plane of the cortex is a critical factor in determining which neurones retain and which lose their pyramidal tract collaterals.

Pieces of fetal rat rostral or occipital cortex labelled with ³H-thymidine (³H-TdR) were transplanted to the rostral cortex of newborn rat pups. Twenty-five days later, the retrogradely transported fluorescent dye Fast blue, was injected into the pyramidal decussation of these animals (see Fig. 1 legend for experimental details). Their brains were examined by fluorescent microscopy 5 days later.

The transplants could not always be recognized confidently by fluorescent microscopy of the unstained sections; however, the presence of ³H-TdR in the donor cells allowed us to identify and delineate accurately the boundaries of the transplants in the autoradiographs under dark-field optics (Fig. 1b). Although the transplants were incorporated within the host cortex to form

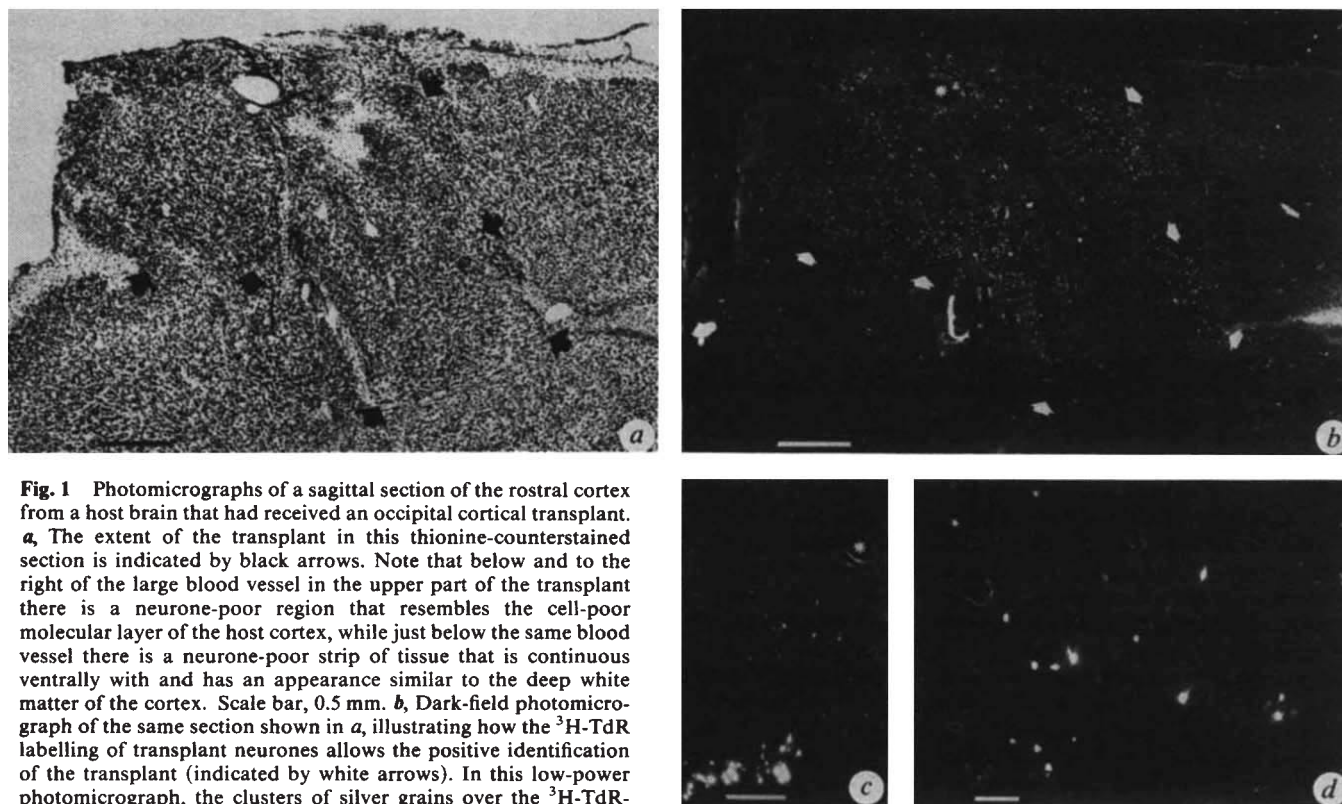


Fig. 1 Photomicrographs of a sagittal section of the rostral cortex from a host brain that had received an occipital cortical transplant. **a**, The extent of the transplant in this thionine-counterstained section is indicated by black arrows. Note that below and to the right of the large blood vessel in the upper part of the transplant there is a neurone-poor region that resembles the cell-poor molecular layer of the host cortex, while just below the same blood vessel there is a neurone-poor strip of tissue that is continuous ventrally with and has an appearance similar to the deep white matter of the cortex. Scale bar, 0.5 mm. **b**, Dark-field photomicrograph of the same section shown in **a**, illustrating how the ^3H -TdR labelling of transplant neurones allows the positive identification of the transplant (indicated by white arrows). In this low-power photomicrograph, the clusters of silver grains over the ^3H -TdR-labelled cells appear as white dots. The asterisk marks the large blood vessel seen in **a**. Scale bar, 0.5 mm. **c**, Photomicrograph taken under fluorescent illumination showing part of the same section illustrated in **a** and **b**. The asterisk is within the same blood vessel referred to in **a** and **b**. Individual and clusters of Fast blue-labelled neurones in the host rostral cortex can be seen in the lower left, while in the centre there are some Fast-blue labelled neurones within the occipital cortical transplant. Scale bar, 0.5 mm. **d**, A higher magnification photomicrograph of the Fast blue-labelled transplant neurones shown in **c**. Scale bar, 0.1 mm.

Methods: Brains were removed from embryonic day 17 Sprague-Dawley rat fetuses taken from pregnant females which, 2 days earlier, had received a single intraperitoneal injection of $[\text{Me-}^3\text{H}]\text{thymidine} (^3\text{H}\text{-TdR}, 10 \mu\text{Ci per g body weight, specific activity } 6.7 \text{ Ci mmol}^{-1}, \text{NEN})$. Small pieces (about $1 \text{ mm} \times 1 \text{ mm}$) of cortex from either the rostral region of cortex ($n=6$) or the occipital region ($n=10$) were dissected out in ice-cold Ham F10 (Gibco) medium before the remainder of the donor brain was fixed in Carnoy's solution for paraffin sectioning and autoradiography²⁴. These pieces of cortex were trimmed further and then placed in a small hole that had been made in the rostral cortex (that is, the frontal region or the anterior parietal region) of the host pups, which had been born within the previous 24 h. The host pups were marked for later identification and returned to their mothers. On postnatal day 25, the host rats were anaesthetized with chloral hydrate and $0.3 \mu\text{l}$ of a 2% suspension of the retrogradely transported dye Fast blue was injected into the pyramidal decussation at the spino-medullary junction. The animals were re-anaesthetized after a 5-day survival and perfused with a 10% buffered formalin solution. The brains were removed, postfixed in 10% sucrose/10% buffered formalin and sagittally sectioned on a freezing microtome. Sections ($30 \mu\text{m}$ thick) were mounted onto gelatin-coated slides, air dried and examined and photographed under fluorescent illumination. The slides were then dipped in NTB-2 emulsion for autoradiography²⁴. After a 4-week exposure, the slides were developed and mounted on coverslips using a non-fluorescent buffered glycerol mountant. The sections were re-examined and photographed with bright, dark-field and fluorescent microscopy, and some sections were subsequently counterstained with thionine.

a more or less continuous tissue, there seemed to be little mixing of host and transplant neurones and the transplant remained a cohesive entity. Occipital cortical transplants and rostral cortical transplants survived equally well in the rostral cortical locale and the two types of transplant were indistinguishable in appearance. Both types of transplant displayed varying degrees of cytoarchitectonic organization. Thionine-stained sections showed clearly that the transplants contained large numbers of healthy neurones, and there were often neurone-poor regions which resembled the normal molecular layer or appeared to be routes taken by fascicles of axons (Fig. 1a).

The injection of Fast blue into the pyramidal decussation resulted in the retrograde labelling of a large number of layer V neurones in the rostral cortex of the host and, as reported by others⁷, in a variable number of Fast blue-labelled neurones within the rostral cortical transplants. Most importantly, a variable but significant number of Fast blue-labelled neurones were recognizable within the occipital cortical transplants (Fig. 1c). At high magnification, the dye-labelled cells in the transplants could be recognized as pyramidal cells, as the proximal parts of their principal dendrites were labelled (Fig. 1d). In the

autoradiographs, a small proportion of the dye-labelled neurones within the transplants was overlain with clusters of silver grains, indicating that these neurones were also labelled with ^3H -TdR (Fig. 2). The Fast blue labelling of these ^3H -TdR-labelled cells was always weak, undoubtedly because the low penetrance of the β -particles emitted by tritium allows the identification only of those ^3H -TdR-containing cells whose nuclei are within the upper few micrometres of the section⁸, and most of these neurones retain relatively little of their Fast blue-containing cytoplasm within the section. However, the occurrence of such doubly labelled neurones is important because it rules out the possibility that the dye-labelled neurones had migrated into the transplants from the host tissue. Note that no dye-labelled neurones were seen in the occipital cortex of the host brains.

We have shown previously that fluorescent dye injections into the pyramidal decussation of rats during the first postnatal week label layer V cells throughout the neocortex, including the occipital cortex, but that similar injections made at 20 days of age fail to label any neurones in the occipital cortex². It was shown subsequently that the magnitude of this transient occipital

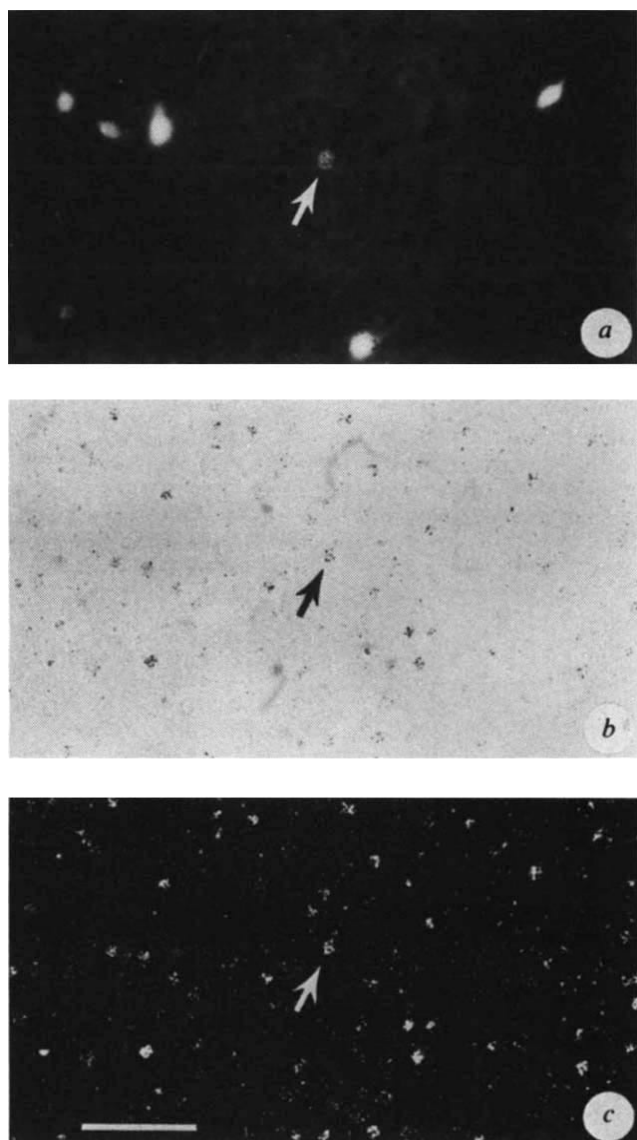


Fig. 2 *a*, Fluorescent photomicrograph of an autoradiograph of some of the same Fast blue-labelled neurones in the occipital cortical transplant shown in Fig. 1*d*. The dye-labelled cell in the centre of the field, which is near the top of the section, has been brought into focus. This neurone is also ^3H -TdR labelled, as indicated by the cluster of silver grains in the emulsion layer above the cell. *b*, *c*, Bright-field and dark-field photomicrographs, respectively, of the same field shown in *a* to illustrate the ^3H -TdR labelling more clearly. The arrow indicates the same doubly labelled neurone in each photomicrograph. Scale bar, 0.1 mm.

cortical component of the pyramidal tract is greatest at about the end of the first postnatal week, after which it steadily declines^{4,6}. Thus, the finding that Fast blue injections into the pyramidal decussation on postnatal day 25 result in labelling of neurones within occipital cortical transplants in the rostral cortex indicates that these neurones (which were placed in the host cortex at a stage when some host pyramidal tract neurones had already extended axons as far as the pyramidal decussation⁹) have not only sent axons through the pyramidal tract, but also have maintained these axons beyond the stage at which such collaterals are normally eliminated. While we cannot be certain that the dye-labelled cells within our occipital cortical transplants are the same neurones which, if left *in situ*, would have given rise to transient pyramidal tract collaterals (it is

difficult to imagine how this could be demonstrated since at the stage of transplant removal pyramidal tract axons have not yet entered the cerebral peduncles⁹), this is the most straightforward interpretation of our observations, and is supported by the fact that some of these neurones are also ^3H -TdR labelled, as it is known that many layer V neurones in the occipital cortex are radioactively labelled in adult rats that have been exposed to a pulse of ^3H -TdR on embryonic day 15 (ref. 10).

Our findings are consistent with observations on developing callosal connections¹¹⁻¹⁷ which indicate that otherwise transient axon collaterals can be stabilized and become permanent following certain experimental manipulations. Our results also demonstrate that this process can be influenced by the position of the neurone within the tangential plane of the cortex. Although we have found that fetal occipital cortical neurones placed in the rostral cortex exhibit one of the projections characteristic of the rostral cortex, we have not yet determined whether other connections characteristic of this region (for example, connections with the appropriate thalamic nuclei) are also established by the transplanted neurones.

Little is known about the factors that contribute to the development of the many striking differences among the various areas of the adult cortex. Our previous finding that during early postnatal development, pyramidal tract neurones, although confined to layer V, are spread virtually throughout the tangential plane of the neocortex suggested that, early in its postnatal development, the neocortex may be more homogeneous than previously recognized². Our present results support this notion and, taken together, our observations suggest: (1) that, at least from the point of view of the connections they establish, different regions of the developing cerebral cortex are not readily distinguishable and in fact may be interchangeable, and (2) that the factors determining at least some of the regional differences in the cerebral cortex are not intrinsic to the various neuronal populations, but are imposed on them by certain extrinsic factors, and that these, in turn, are a function of the topographic location of the cortical region.

Finally, our results do not imply that there are no intrinsic differences among cortical neurones. Laminar differences in the efferent connections of cortical neurones seem to emerge quite early², and these connectional differences seem to be maintained by the neurones even when their laminar position is grossly altered¹⁸⁻²³.

We thank Mary Ann Lawrence and Kris Trulock for technical assistance, Pat Thomas for secretarial help, and Dr W. M. Cowan for his comments on an earlier version of this manuscript. This work was supported by NIH grants NS-18506 and EY-03653.

Received 17 September; accepted 22 October 1984.

1. Wise, S. P., Murray, E. A. & Coulter, J. D. *Neuroscience* **4**, 65-78 (1979).
2. Stanfield, B. B., O'Leary, D. D. M. & Fricks, C. *Nature* **298**, 371-373 (1982).
3. Bates, C. A. & Killackey, H. P. *Devl Brain Res.* **13**, 265-273 (1984).
4. Stanfield, B. B. & O'Leary, D. D. M. *Soc. Neurosci. Abstr.* **8**, 438 (1982).
5. Stanfield, B. B. & O'Leary, D. D. M. *Soc. Neurosci. Abstr.* **9**, 375 (1983).
6. Stanfield, B. B. & O'Leary, D. D. M. *J. comp. Neurol.* (submitted).
7. Floeter, M. K. & Jones, E. G. *J. Neurosci.* **4**, 141-150 (1984).
8. Rogers, A. W. *Techniques of Autoradiography* (Elsevier, Amsterdam, 1973).
9. Schreyer, D. J. & Jones, E. G. *Neuroscience* **7**, 1837-1853 (1982).
10. Bruckner, G., Mares, V. & Biesold, D. *J. comp. Neurol.* **166**, 245-256 (1976).
11. Lund, R. D., Mitchell, D. E. & Henry, G. H. *Brain Res.* **144**, 169-172 (1978).
12. Innocenti, G. M. & Frost, D. O. *Nature* **280**, 231-234 (1979); *Exptl Brain Res.* **39**, 365-375 (1980).
13. Rhoades, R. W. & Dellacrose, D. D. *Brain Res.* **246**, 146-149 (1982).
14. Cusick, C. G. & Lund, R. D. *J. comp. Neurol.* **212**, 385-398 (1982).
15. Rothblat, L. A. & Hayes, L. L. *Brain Res.* **246**, 146-149 (1982).
16. Berman, N. E. & Payne, B. R. *Brain Res.* **274**, 201-212 (1983).
17. Lund, R. D., Chang, F.-L. F. & Land, P. W. *Devl Brain Res.* **14**, 139-142 (1984).
18. Caviness, V. S. Jr & Yorke, C. H. *J. comp. Neurol.* **170**, 449-460 (1976).
19. Caviness, V. S. Jr & Rakic, P. A. *Rev. Neurosci.* **1**, 279-326 (1978).
20. Lemmon, V. & Pearlman, A. L. *J. Neurosci.* **1**, 83-93 (1981).
21. Simmons, P. A., Lemmon, V. & Pearlman, A. L. *J. comp. Neurol.* **211**, 295-308 (1982).
22. Terashima, T., Inoue, K., Inoue, Y., Mikoshiba, K. & Tsukada, Y. *J. comp. Neurol.* **218**, 314-326 (1983).
23. Jensen, K. F. & Killackey, H. P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 964-968 (1984).
24. Cowan, W. M., Gottlieb, D. I., Hendrickson, A. E., Price, J. L. & Woolsey, T. A. *Brain Res.* **37**, 21-51 (1972).

A zone of non-proliferating cells at a lineage restriction boundary in *Drosophila*

David A. O'Brochta & Peter J. Bryant

Developmental Biology Center and Department of Developmental and Cell Biology, University of California at Irvine, Irvine, California 92717, USA

The use of X-ray-induced mitotic recombination to genetically mark individual cells and their descendants during development has led to the discovery of lineage restriction boundaries in *Drosophila* imaginal disks, dividing the disks into areas called compartments^{1,2}. Clones of cells initiated after a given developmental stage are unable to grow across these boundaries, even if provided with a growth rate advantage over the remaining cells. It has been suggested that cells within compartments are distinguished by the differential activation of selector genes³ and that the lineage restrictions are maintained by adhesivity differences between the cells in different compartments⁴, but other mechanisms have not been ruled out. Recently a discrete population of cells with unusual permeability properties has been described along an intersegmental lineage restriction boundary in *Oncopeltus*⁵, suggesting that a lineage restriction could be maintained by a zone of cells which present a barrier to clone growth. Here we demonstrate by autoradiography the presence of a narrow zone of non-proliferating cells (ZNC) coincident with the presumptive wing margin in the *Drosophila* wing disk, and suggest that this could account for the observed lineage restriction between presumptive dorsal and ventral surfaces of the wing. As the anterior/posterior compartment boundary does not coincide with a ZNC, the results indicate that different lineage boundaries may be maintained by different mechanisms.

We analysed autoradiographs of serially sectioned imaginal wing disks that were labelled *in vitro* with tritiated thymidine after explantation from larvae at different stages. Late third-instar wing disks incorporated label in the columnar epithelium, peripodial membrane and adepithelial cells, but they invariably showed a narrow zone of cells across the central portion of the wing pouch centred on the basal groove⁶ in which the labelling index (labelled nuclei/total nuclei) was greatly reduced (Figs 1, 2). This region of reduced labelling, which we interpret as a zone of non-proliferating cells (ZNC), was also seen in disks labelled *in situ* and in disks cut in half longitudinally before incubation to ensure access of label into the recesses of the wing pouch. It is therefore not an artefact of labelling *in vitro* or non-uniform access to tritiated thymidine.

In the late third instar (120 h after egg laying; AEL), the ZNC is approximately 30 μm or 6–10 cells wide. High-resolution analysis of selected disks shows that there is no labelling in the central 10 μm of the zone, with gradients of labelling index approximately 10 μm wide on either side. Serial section reconstructions demonstrate that the ZNC coincides with the position of the presumptive wing margin and the position of the dorsal/ventral lineage restriction boundary across the wing pouch (Fig. 3). A depressed labelling index, quantitatively less pronounced than that seen in the late third instar, was also detected in the presumptive wing margin region of wing disks from larvae 96 and 108 h AEL, corresponding to middle and mid-late third instar, respectively (Fig. 2).

No other region of the imaginal wing disk has growth properties like that of the presumptive wing margin, although we have found other evidence of non-uniform growth. The base of each fold of the late third-instar wing disk shows a consistently low labelling index relative to nearby regions of the epithelium. This trend generally also holds for mid-third-instar (96 h AEL) wing disks, although the resolution of the analysis for this stage is low because of the small size of the disks. Late third-instar

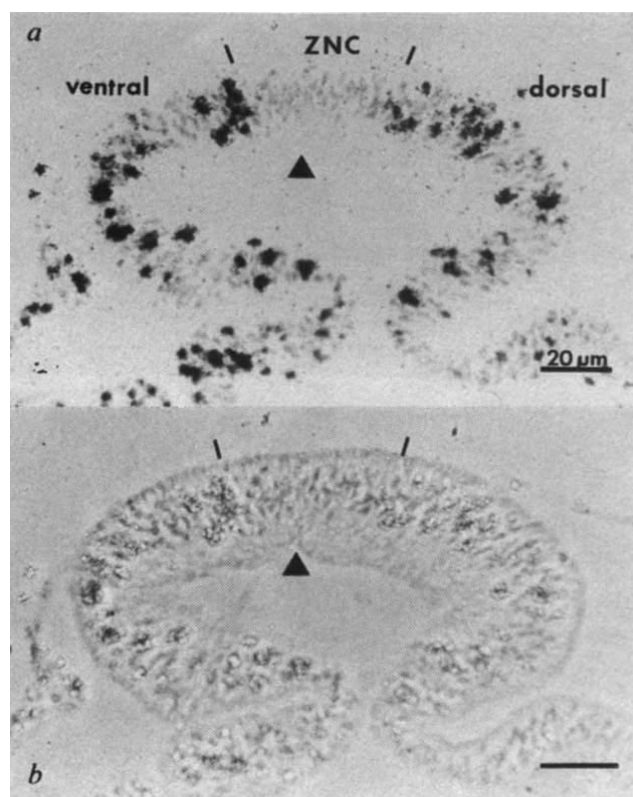
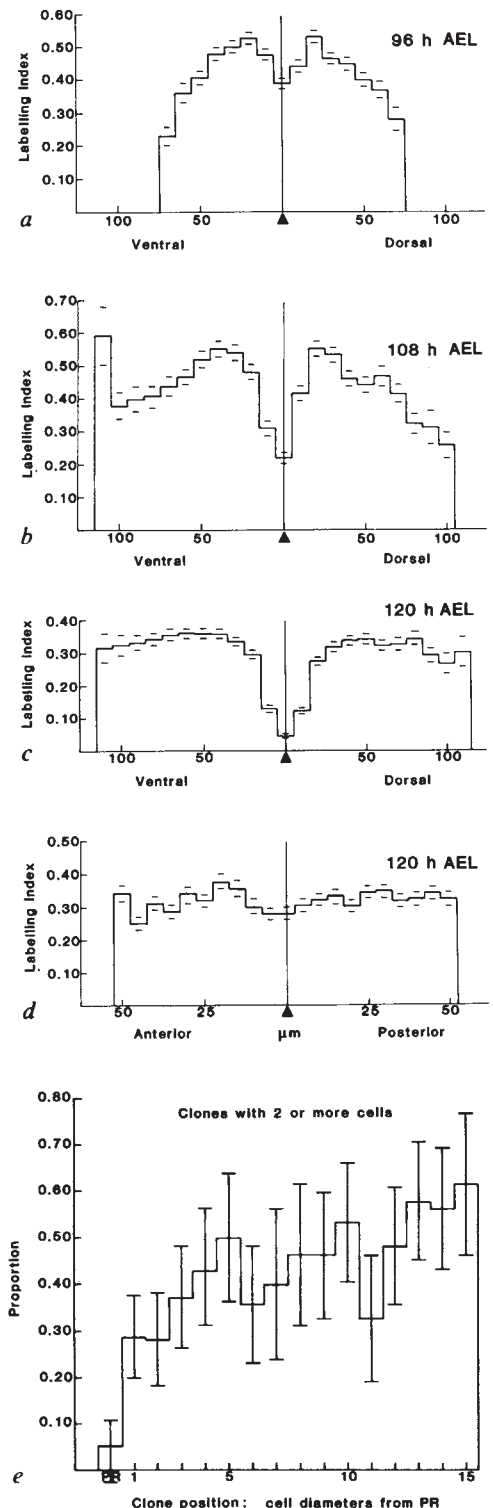


Fig. 1 Autoradiograph (a) and phase-contrast micrograph (b) of a late third-instar wild-type wing pouch after being labelled *in vitro* with tritiated thymidine and sectioned perpendicular to the dorsal/ventral lineage restriction boundary. The ZNC is clearly visible (a) and includes the cells on either side of the basal groove (b, arrow).

Methods: Imaginal disks were dissected in *Drosophila* Ringer's solution²⁴ and transferred individually to 50 μl of Ringer's solution containing 0.1 $\mu\text{Ci } \mu\text{l}^{-1}$ of ^3H -thymidine ([Me- ^3H]thymidine, specific activity 77–83 Ci mmol⁻¹, NEN). Disks were incubated in a covered chamber at room temperature (23–26 °C) for 30 min, rinsed in 0.5 ml Ringer's solution and then transferred to 50 μl of Ringer's containing 0.25 M sucrose for 15 min. (The second washing results in the nuclei appearing more distinct, facilitating identification of labelled and unlabelled nuclei.) The disks were then fixed in 0.5 ml of Kahle's²⁵ fixative containing Fast green (Fisher, CI 42053) for ~45 min. The tissue was either stored in 70% ethanol at 4 °C before processing or processed immediately and serially sectioned longitudinally at 5 μm . De-waxed and hydrated sections were hydrolysed for 10 min in 6 N HCl at room temperature, rinsed in water and stained in Schiff's reagent²⁵ for 75 min at room temperature in the dark. (Schiff's reagent²⁵ was prepared with twice the recommended concentration of $\text{Na}_2\text{S}_2\text{O}_5$.) The stained slides were washed twice in 2.4×10^{-2} M $\text{Na}_2\text{S}_2\text{O}_5$ in 4.5×10^{-2} N HCl for 2 min followed by a 30-min wash in water. Autoradiographs were prepared by dipping the dried stained slides in a 1:1 mixture of distilled water and Kodak NTB-2 nuclear track emulsion at 40 °C, drying them vertically for 1 h and placing them in a light-tight box containing desiccant for 12 h at –70 °C. The slides were removed from the freezer 30 min before developing and developed in a Kodak D19 developer for 3 min at 20 °C, fixed in Kodak Rapid Fix for 1.5 min, washed in water for 30 min, and coverslips were put in place.

imaginal wing disks sectioned transversely did not have an obvious zone with a reduced labelling index along the anterior/posterior compartment boundary, and a quantitative analysis of the autoradiographs used in our dorsal/ventral analysis confirms this conclusion (Fig. 2). All other imaginal disks analysed had a non-uniform distribution of labelled cells, but a ZNC was not detected in the capitellum pouch of the haltere disk. This is consistent with the finding that fast-growing clones in the haltere disk do not seem to recognize a dorsal/ventral boundary⁷; the fact that clones which are transformed into wing

Fig. 2 *a-c*, Profile of average labelling index in the longitudinal dimension (perpendicular to the dorsal/ventral lineage restriction boundary, see Fig. 3) in the wing disk. *a*, 96 h AEL; *b*, 108 h AEL; *c*, 120 h AEL, mean $\pm$ standard error (bars). The transverse folds, the wing pouch and the basal groove along the centre of the wing pouch extend transversely across most of the width of the wing disk, so longitudinal sections of the central 100–125 μ m of the late third-instar wing disk appear uniform in shape. Taking advantage of the similarity of sections and the morphological features mentioned above, the data were averaged across sections. Camera lucida drawings of each section were made in which each nucleus was scored as either labelled or unlabelled. Using the basal groove as a reference point, 10- μ m sectors were drawn on each section and the number of labelled and unlabelled nuclei recorded. Data from corresponding sectors from successive sections were combined and the mean labelling index computed (mean no. of labelled nuclei in sector/mean no. of nuclei in sector). In cases where the basal groove was not visible, the centre of the wing pouch was used as a reference point. Approximately 200 sections (10 disks) were included in our analysis of 96-h wing disks, 100 sections (5 disks) from 108-h wing disks and 230 sections (10 disks) from 120-h wing disks. *d*, Profile of average labelling index in the transverse direction (perpendicular to the anterior/posterior lineage restriction boundary, see Fig. 3) 120 h AEL, mean $\pm$ s.e. Average labelling index data were generated as in *a-c*, but in this case the mean labelling index for the wing pouch region of each longitudinal section was calculated by combining the data from all sectors in the wing pouch except the central three (corresponding to the ZNC). Taking the middle of the wing pouch as an approximation of the border between the presumptive anterior and posterior cells, the labelling index profile for an average transverse section of a 120-h wing pouch was calculated. *e*, Proportion of $y^-; mwh$ clones that contain two or more cells (clones of two cells or more/total clones) as a function of distance from the posterior wing margin. Clones were induced by γ -irradiation 0–8 h before pupariation. Ninety-five % confidence limits are shown. Fifty-four wings were analysed. Uncrowded third instar *Df(1)sc⁸, w^a; Dp(1;3)sc¹⁴/mwh* larvae¹⁵ cultured at 25 °C were irradiated with 1,000 rad of γ rays at 1,000 rad per min using an Isomedix M38-2 ¹³⁷Cs source. In this genotype the y^+ locus is translocated to the left arm of chromosome 3, which is also the location of *mwh*⁺. The *Df(1)sc⁸* X chromosome is deficient for the y^+ locus. Mitotic recombination in the left arm of chromosome 3 results in the elimination of y^+ and *mwh*⁺ from one daughter cell, giving clones marked with y^- and *mwh* (ref. 19). Pupae were collected at 8-h intervals after irradiation and wings from the resulting male flies were mounted between cover slips. Only the posterior wing margin region was scored for the presence of $y^-; mwh$ clones. The posterior row elements are socketless hairs and can be scored for y^- , while the trichomes within the wing blade can be scored for *mwh*. This enabled us to score every cell comprising the posterior wing margin. The location of a marked cell or clone of marked cells was measured by counting the number of trichomes between the marked cell closest to the margin and the margin itself. Clones produced 0–8 h before pupation can be found in any position including the wing margin; however, multiple cell clones are more likely to be found distant from the margin.



by homoeotic mutations appear to recognize a dorsal/ventral boundary in the haltere⁸ does not imply that a boundary is present in the haltere disk itself. The cells which form the most distal segments of all leg disks and the antenna portion of the eye-antenna disk had a greatly reduced labelling index. A ZNC was not detected in previous autoradiographic analyses of *Drosophila* imaginal wing disks^{9,10}, but this can be largely attributed to the methods used. While we have occasionally seen evidence of the ZNC using whole-mount autoradiography⁹, we have found that this method gives inconsistent results. The zone is probably too narrow to have been detected as a difference in labelling index between dissected fragments¹⁰.

Although this is the first direct demonstration that the presumptive wing margin region has unusual growth properties,

other indirect evidence is consistent with these findings. Mitotic recombination data indicate that precursor cells for the medial bristles of the triple row portion of the wing margin cease proliferative divisions in the middle of the third instar, whereas the remainder of the disk continues growing¹¹. However, the ZNC is not composed solely of bristle mother cells; cells responsible for producing the posterior row elements of the wing, which are non-innervated socketless hairs, are also included in the ZNC.

We have found that clones induced within the last 8 h of the third instar are as likely to be found in the posterior row as in any of the first 15 rows of trichomes counting from the margin (mean no. of clones/posterior row/wing = 1.00 ± 0.18 , mean no. of clones/rows 1–15/wing = 0.88 ± 0.05 ; there is no significant

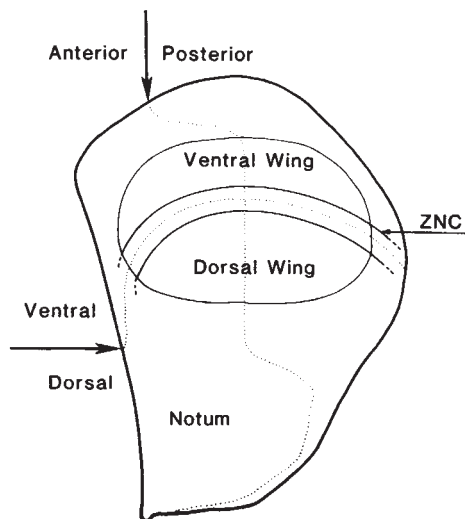


Fig. 3 A summary diagram based on serial reconstructions showing the location of the ZNC in relation to anterior/posterior and dorsal/ventral lineage restriction boundaries (dotted lines)^{26,27}. The location of the presumptive wing structures is based on the fate map of Bryant²⁸.

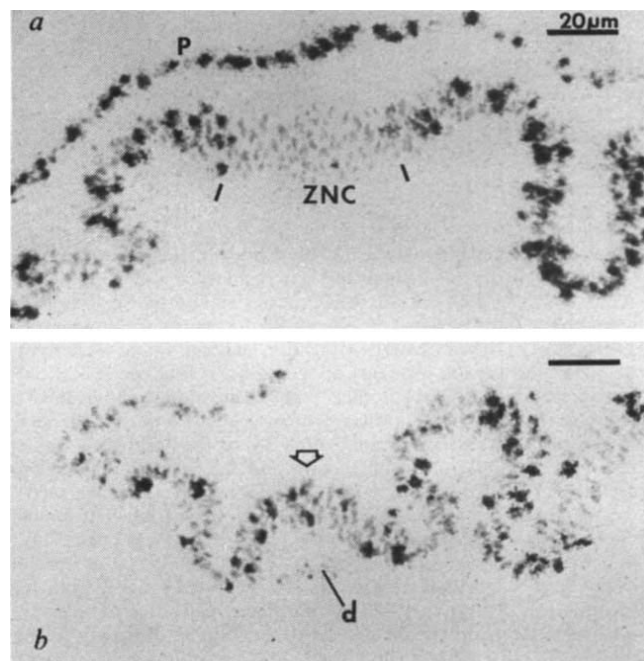


Fig. 4 Autoradiographs of imaginal wing disk sections from mutants homozygous for *lethal (2) giant discs* (ref. 14) *a* and *vestigial* (ref. 15) *b* raised at 25°C. In both cases the disks were labelled *in vitro* as previously described and sectioned longitudinally. Only the wing pouch region is shown. *a*, The *l(2)gd* wing disk was from a 10-day-old larva and has a ZNC with well-defined borders and a high frequency of labelling in the remaining epithelium and peripodial membrane (P). *b*, The *vg* wing disk, which shows evidence of cell death in the wing pouch¹⁶ (d), was from a late third-instar larva and lacks a ZNC (arrow).

difference between these means at $P=0.05$; 36 wings were analysed). However, only 5% of the clones located within the posterior row contained more than one cell, while 50% of those located distant from the margin contained more than one cell (Fig. 2). This is consistent with the idea that the prospective margin is composed of non-dividing cells at this time. The fact that any marked cells are found along the margin is attributed to the fact that a high proportion of cells in the late third-instar wing disk have a 4C DNA content^{12,13} and can therefore divide without further DNA replication.

The development of the ZNC is a disk-autonomous phenomenon. Of 50 wing disks collected from animals in the process of moulting from the second to the third instar and cultured *in vivo* for 2, 3 or 4 days at 25°C in the abdomens of adult females, 16 (32%) showed a ZNC at the appropriate location. The remainder probably failed to reach the stage of development at which the zone becomes qualitatively obvious. Growth termination in the remainder of the disk also occurs autonomously in transplanted disks, at a later stage in the culture period (data not shown). The two growth termination events are genetically distinguishable, as shown by the fact that the overgrowing imaginal wing disks from the mutant *lethal (2) giant discs* (ref. 14) develop a ZNC despite the continued proliferation of the remaining cells of the disk (Fig. 4a). The ZNC in these disks is wider than in wild type and shows a more abrupt transition between growing and non-growing regions.

The ZNC represents a discrete population of cells with special growth properties, rather than a response of cells to the confrontation between presumptive dorsal and ventral regions of the wing disk. Late third-instar imaginal wing disks from three mutants lacking the wing margin (*vestigial*, *Ultravestigial* and *nubbin*¹⁵) also lack the ZNC, despite the direct apposition of presumptive dorsal and ventral cells (Fig. 4b)¹⁶.

The presence of the ZNC can account for the observed lineage restrictions between presumptive dorsal and ventral wing regions, if cells that enter the zone as a result of clone growth are induced to reduce their proliferation rate (at the edges of the zone) or stop dividing entirely (at the centre of the zone). According to this interpretation, clones would recognize two parallel restriction lines if they were induced outside the ZNC and grew up to the ZNC late in the third instar. Exactly this result has been reported in clonal studies of the undifferentiated wing disk, and the distance between the observed restriction lines corresponds closely with the width of the ZNC¹⁷.

The absence of two restriction boundaries in the adult wing could result from loss of the ZNC by cell death during pupal

development. However, this seems unlikely because cell death has not been observed in this region during pupal development¹⁸, and mitotic recombination evidence (Fig. 3 and ref. 9) suggests that the ZNC contributes to the adult wing. Alternatively, the results could be explained by the gradual recruitment of formerly mitotic cells into the zone during the late third instar. When the ZNC is established, it is composed of fewer cells than it will ultimately include in the late third instar; in fact, in the early third instar it could consist of a single row of non-dividing cells. Therefore, formerly dividing cells must be recruited into the zone to allow for its enlargement. Such a mechanism would result in a discrete ZNC being detected in autoradiographic experiments, but it is also consistent with one, rather than two, restriction boundaries in the adult cuticle.

We suggest that the dorsal/ventral lineage restriction^{2,19} is a secondary consequence of the presence of a ZNC in the wing disk. Our data indicate the presence of a ZNC at the middle of the third instar and we cannot exclude the possibility that a narrow ZNC is present earlier. In previous clonal analyses utilizing the *Minute* technique to provide marked clones with a growth-rate advantage²⁰, the latest time at which clones were induced and subsequently included both dorsal and ventral cells was 96–120 h² and 72 ± 2 h AEL²¹. In the *Minute* mutant used in both of these studies, *M(3)i*²⁵, the third larval instar begins at approximately 75 h AEL and ends at 155 h AEL²². Therefore, the dorsal/ventral lineage boundary appears to be established during the third instar, coinciding with the appearance of non-proliferating cells at the appropriate location.

Our finding of a ZNC at the dorsal/ventral boundary but not at the anterior/posterior boundary adds to the growing body of evidence suggesting that different cellular mechanisms are involved in establishing and maintaining these two boundaries in the wing disk. The anterior/posterior boundary is established before wing disk formation, but the dorsal/ventral boundary is established much later within the disk itself^{1,2}. The dorsal/ven-

tral boundary coincides with the wing margin, but the anterior/posterior boundary does not coincide with any anatomical feature^{1,2}. Mutations are known which affect either the anterior or posterior compartment, but analogous mutations affecting the dorsal or ventral compartment have not been reported³. Certain monoclonal antibodies show preferential binding to dorsal or ventral compartments, but corresponding distinctions between anterior and posterior compartments have not been found²³. We conclude that the apparently sequential subdivision of a tissue into areas called compartments may result from a series of fundamentally different developmental events, rather than the repeated operation of the same type of event.

This investigation was supported by NIH research grant HD06082 and NIH National Research Service Award HD07029. We thank Dr Danny Brower for helpful discussions and comments.

Received 9 April; accepted 20 October 1984.

1. Garcia-Bellido, A., Ripoll, P. & Morata, G. *Nature* **245**, 251–253 (1973).
2. Garcia-Bellido, A., Ripoll, P. & Morata, G. *Dev. Biol.* **48**, 132–147 (1976).
3. Garcia-Bellido, A. *Ciba Fdn Symp.* **29**, 161–182 (1975).
4. Morata, G. & Lawrence, P. A. *Nature* **255**, 614–617 (1975).
5. Blennerhassett, M. G. & Caveney, S. *Nature* **309**, 361–364 (1984).
6. Brower, D. L., Smith, R. J. & Wilcox, M. *J. Embryol. exp. Morph.* **67**, 137–151 (1982).
7. Hayes, P. thesis, Univ. Alberta (1982).
8. Morata, G. & Garcia-Bellido, A. *Wilhelm Roux Arch. dev. Biol.* **179**, 125–143 (1976).
9. Dale, L. & Bownes, M. *Wilhelm Roux Arch. dev. Biol.* **189**, 91–96 (1980).
10. Adler, P. N. & MacQueen, M. *Dev. Biol.* **103**, 28–37 (1984).
11. Garcia-Bellido, A. & Merriam, J. R. *Dev. Biol.* **24**, 61–87 (1971).
12. Fain, M. J. & Stevens, B. *Dev. Biol.* **92**, 247–258 (1982).
13. Graves, B. J. & Schubiger, G. *Dev. Biol.* **93**, 104–110 (1982).
14. Bryant, P. J. & Schubiger, G. *Dev. Biol.* **24**, 233–263 (1971).
15. Lindsley, D. L. & Grell, E. H. *Carnegie Inst. Wash. Publ. no. 627* (1968).
16. O'Brochta, D. & Bryant, P. J. *Wilhelm Roux Arch. dev. Biol.* **192**, 285–294 (1983).
17. Kuhn, D. T., Fogerty, S. C., Eskens, A. C. & Sprey, Th. E. *Dev. Biol.* **95**, 399–413 (1983).
18. Waddington, C. H. *J. Genet.* **41**, 75–139 (1940).
19. Bryant, P. J. *Dev. Biol.* **22**, 389–411 (1970).
20. Morata, G. & Ripoll, P. *Dev. Biol.* **42**, 211–221 (1975).
21. Morata, G. & Lawrence, P. A. *Dev. Biol.* **70**, 355–371 (1979).
22. Steiner, E. *Wilhelm Roux Arch. dev. Biol.* **180**, 9–30 (1976).
23. Brower, D. L., Wilcox, M., Piovant, M., Smith, R. J. & Reger, L. A. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
24. Schubiger, G. *Dev. Biol.* **28**, 277–295 (1971).
25. Thompson, S. N. *Selected Histochemical and Histopathological Methods*, 353, 1039–1039 (Thomas, Springfield, Illinois, 1966).
26. Wilcox, M., Brower, D. L. & Smith, R. J. *Cell* **25**, 159–164 (1981).
27. Brower, D. L., Lawrence, P. A. & Wilcox, M. *Dev. Biol.* **86**, 448–455 (1981).
28. Bryant, P. J. *J. exp. Zool.* **193**, 49–78 (1975).

Inositol 1,4,5-trisphosphate mimics muscarinic response in *Xenopus* oocytes

Y. Oron, N. Dascal, E. Nadler & M. Lupu

Department of Physiology and Pharmacology, Sackler Faculty of Medicine, Tel Aviv University, Ramat Aviv 69978, Israel

The enhanced metabolism of phosphoinositides, which is associated with a wide variety of stimuli and physiological responses, has been studied intensively. Berridge and his collaborators^{1–3} demonstrated that the first measurable reaction following cell membrane receptor activation is a rapid hydrolysis of phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P₂), and that the product of this reaction, inositol 1,4,5-trisphosphate (Ins(1,4,5)P₃), could cause a release of non-mitochondrial calcium. These findings have been verified in other systems^{4–6}. Although the relationship between the hydrolysis of PtdIns(4,5)P₂ and the mobilization of intracellular calcium was clearly demonstrated, the direct link between Ins(1,4,5)P₃ production and the physiological response was only implied. We have investigated the possibility that the intracellular release of Ins(1,4,5)P₃ mediates the muscarinic-cholinergic response in *Xenopus* oocytes, and we show here that intracellularly injected Ins(1,4,5)P₃ mimics the muscarinic depolarizing chloride current in *Xenopus* oocytes. This is the first demonstration of a direct link between phosphoinositides metabolism and a neurotransmitter-induced physiological response.

Bath application of muscarinic agonists to *Xenopus laevis* oocytes results in a complex response consisting of two consecutive inward currents carried by chloride ions (D₁ and D₂), a

Table 1 Effect of ACh on phosphoinositide breakdown products

% Increase in radiolabel (mean ± s.e.m., n = 6)				
Inositol	GPI	Ins(1)P	Ins(1,4)P ₂	Ins(1,4,5)P ₃
–5 ± 11	18 ± 4	37 ± 7	74 ± 11	85 ± 17

Stage 6 *Xenopus* oocytes were separated from the enclosing follicular cells (denuded) by treatment with collagenase⁹. The cell suspension was incubated overnight at room temperature with gentle shaking, in normal Ringer's solution, in the presence of 20–50 µCi ml^{–1} ³H-inositol. The cells were rinsed, resuspended in normal Ringer's solution and incubated for 120 s in the presence of 0.1 mM ACh. The assay was terminated by homogenization in 5% trichloroacetic acid (TCA). Proteins were removed by centrifugation and the TCA extracted with ether. The extracts were diluted and the products of phosphoinositide breakdown were separated in Dowex AG21K columns (formate form), essentially as described by Downes and Michell¹⁴. Each experimental system contained 50–100 oocytes. The results are expressed as a mean % increase in radioactivity over control (no additions) ± s.e.m. The mean distribution of radioactivity between the various ³H-inositol-labelled compounds was as follows (c.p.m. ± s.e.m., 17 experiments in untreated oocytes, approximate counting efficiency 20%): free intracellular inositol, 43,000 ± 9,000; glycerylphosphorylinositol (GPI), 320 ± 30; Ins(1)P, 930 ± 150; Ins(1,4)P₂, 120 ± 20; Ins(1,4,5)P₃, 220 ± 50. Membrane associated counts (TCA precipitate) were 9,000 ± 1,500 c.p.m. (mean ± s.e.m., eight experiments). Commercial ³H-inositol contained impurities which eluted from anion exchange columns at points corresponding to various inositol sugar phosphates (particularly Ins(1)P). These impurities probably represent negatively charged labelled compounds, as they could be completely removed by two successive passages through a Dowex AG21K column and would not be expected to be taken up by the cell. In several experiments we used column-purified ³H-inositol and obtained elution patterns identical to those obtained with unpurified tracer.

prolonged outward current carried by potassium ions (H) and superimposed large chloride current fluctuations (F)^{7–10}. D₁, D₂ and F have been shown to be calcium dependent^{9,10}. Moreover, we have found that acetylcholine (ACh) releases oocyte-bound calcium, and that intracellular injection of calcium ions results in a chloride inward current similar to the muscarinic D₁, D₂ and F responses (B. Gillo, N.D. and Y. Lass, unpublished results). Little is known about the physiological role of muscarinic receptors in the amphibian oocyte. ACh has been measured in the *Xenopus* oocyte¹¹ and the oocytes have measurable acetylcholinesterase activity¹². We have reported recently muscarinic promotion of progesterone-induced oocyte maturation¹³; we have characterized this response pharmacologically and have shown that ACh decreases cyclic AMP levels in this cell (N.D., R. Yekuel and Y.O., unpublished). We speculate that the activation of muscarinic receptors modulates the process of oocyte maturation.

The activation of muscarinic receptors may result in the increased hydrolysis of PtdIns(4,5)P₂; the resulting increase in the concentration of Ins(1,4,5)P₃ may cause the mobilization of intracellular calcium, analogous to the mobilization of calcium by Ins(1,4,5)P₃ in other cells. That, in turn, activates membrane chloride channels. To verify this hypothesis, we first demonstrated that cholinergic stimulation of *Xenopus* oocytes resulted in the hydrolysis of PtdIns(4,5)P₂. When oocytes pre-labelled with ³H-inositol were challenged with cholinergic agonists, there was a significant increase in the accumulation of phosphoinositide breakdown products: inositol 1-phosphate (Ins(1)P), inositol 1,4-bisphosphate (Ins(1,4)P₂), and Ins(1,4,5)P₃ (Table 1). This effect was blocked by atropine (0.1–1 µM, added 10 min before ACh, not shown), and could be observed 30–120 s after stimulation with the cholinergic agonist.

The injection of Ins(1,4,5)P₃ into oocytes produced two consecutive inward currents and variable superimposed current fluctuations (Fig. 1) similar to the chloride current responses seen on ACh application (Fig. 1b). Although in this particular denuded oocyte ACh produced only the D₂ response, we could demonstrate that the injection of Ins(1,4,5)P₃ does not require the involvement of the surrounding follicular cell layer. Both

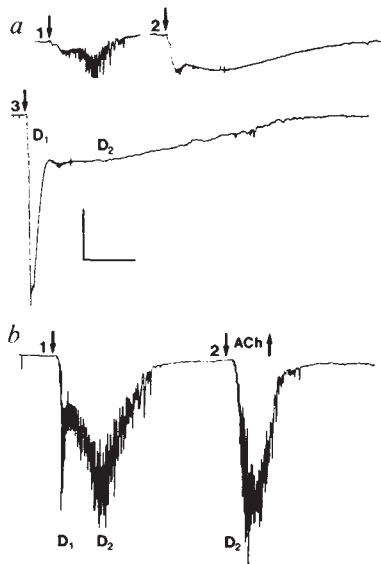


Fig. 1 *a*, Response of a denuded oocyte to pressure injection of increasing doses of Ins(1,4,5) P_3 ; arrow 1, 0.5 pmol; arrow 2, 1.25 pmol; arrow 3, 2 pmol. Calibration: vertical bar, 100 nA; horizontal bar, 6 min. *b*, Response of a single denuded oocyte to iontophoretic injection of Ins(1,4,5) P_3 (arrow 1) and ACh (10 μ M) application (arrow 2). Calibration: vertical bar, 50 nA; horizontal bar, 6 min.

Methods. *a*, Oocytes were placed in a 1-ml bath constantly perfused with normal Ringer's solution and were voltage clamped using a conventional voltage-clamp technique (usually at -50 or -60 mV). A third micropipette filled with 0.1–10 mM Ins(1,4,5) P_3 was used for pressure injections. Ins(1)P, Ins(1,4) P_2 and Ins(1,4,5) P_3 were obtained from rabbit erythrocyte ghosts and separated on a Dowex AG21K formate column, essentially as described by Downes *et al.*¹⁵, except that the LiCl stage was omitted and the formate removed by repeated, prolonged lyophilization. Concentrations were calculated from total phosphorus determination according to Bartlett¹⁶. *b*, Ins(1,4,5) P_3 was injected at 2.4 μ A current for 8 s (maximal amount injected 67 pmol, assuming that all the current was carried by Ins(1,4,5) P_3 ions and that the net charge per molecule was -6). In this particular cell the application of ACh resulted in D_2 and F currents only. This is common in denuded oocytes, which lose some of the components of the complete response. However, we observed a complete response (D_1 , D_2 and F) both to ACh application and to Ins(1,4,5) P_3 injection in many intact (follicle enclosed) oocytes in this series of experiments.

D_1 -like and D_2 -like responses to Ins(1,4,5) P_3 injection were accompanied by an increase in membrane conductance (Fig. 2). The reversal potentials (V_r) were -18.6 ± 2.0 mV (mean $\pm$ s.e.m., $n=5$) for the D_1 -like response and -23.8 ± 2.0 mV (mean $\pm$ s.e.m., $n=8$) for the D_2 -like response. The more negative V_r of the D_2 -like response implies a minor component of potassium current. These values are close to the equilibrium potential of chloride ions in *Xenopus* oocytes (about -22 mV, see ref. 9). In two oocytes, measurement of the reversal potentials of D_1 - and D_2 -like responses to Ins(1,4,5) P_3 injection and a D_1 response to ACh application showed them to be virtually identical (within ± 2 mV, see Fig. 2).

The response to Ins(1,4,5) P_3 was observed in all 17 cells tested (7 frogs, 15 follicles, 2 denuded oocytes). Both pressure and iontophoretic injections were effective (Fig. 1). Iontophoretic injection with reverse polarity (cation iontophoresis) was ineffective (not shown), indicating that the effect was not due to the injection of contaminating calcium ions.

The depolarizing response to Ins(1,4,5) P_3 injection was dose dependent (Fig. 1*a*), with an apparent threshold for injected Ins(1,4,5) P_3 of 0.02–0.1 pmol (in follicles) and 0.25 pmol (in denuded cells) per oocyte, corresponding to a nominal intracellular concentration of 0.02–0.2 μ M. Acid hydrolysis of Ins(1,4,5) P_3 (6 M HCl, 2.5 h at 100 $^{\circ}$ C) resulted in $\sim 30\%$ hydrolysis of organic phosphorus (corresponding to one phosphate

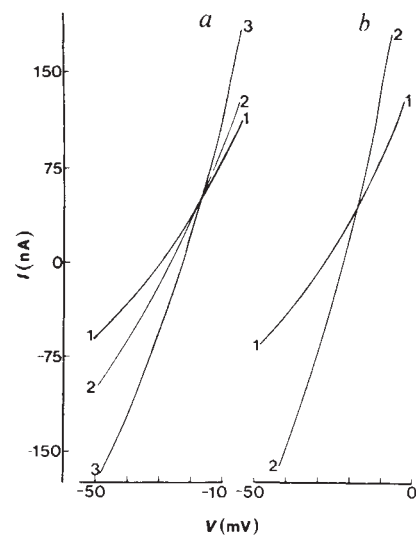


Fig. 2 Current-voltage (I - V) characteristics of the response of a single follicle in Ins(1,4,5) P_3 injection (*a*) and of the muscarinic D_1 response (*b*). The I - V curves were recorded automatically with an X - Y plotter, following a ramp-like change in the holding potential which lasted 0.8–1.0 s (ref. 9). This method has been tested carefully against manual adjustment of holding potential and the reversal potentials were found to be identical when determined by both methods⁹. The reversal potential (V_r) is the point of intersection between the I - V curves of the cell (at rest and at the plateau of the response)⁹. *a*: Trace 1, I - V curve of the resting membrane; trace 2, the I - V curve at the time of the D_2 -like response plateau ($V_r = -17$ mV); trace 3, I - V curve at the time of the D_1 -like response plateau ($V_r = -17$ mV). *b*: Trace 1, I - V curve of the resting membrane; trace 2, I - V curve at the time of the muscarinic D_1 response plateau ($V_r = -18$ mV). ACh was 2 μ M. The I - V curve of the muscarinic D_2 response was not determined in this experiment and can be found in refs 9 and 10.

moiety) and $\sim 80\%$ loss of the response (four cells; unhydrolysed Ins(1,4,5)P was also tested in two of these cells). The injection of ammonium formate (a possible contaminant of the tested compounds) at up to 56 pmol per oocyte produced no response. The injection of Ins(1)P (threshold 1.5–2.0 pmol per oocyte), however, produced a marked depolarizing current, while Ins(1,4) P_2 produced inconsistent responses: either outward or inward currents, associated with both increases and decreases in membrane conductance, were observed. The reversal potentials of the inward current responses to Ins(1)P and Ins(1,4) P_2 injections were all positive— $+10.5 \pm 6.2$ mV, (mean $\pm$ s.e.m., $n=6$) for Ins(1)P in six cells tested and $> +14$ mV for Ins(1,4) P_2 in three out of six cells tested in which the inward current was associated with increased membrane conductance. This reversal potential indicated that the responses to Ins(1)P and Ins(1,4) P_2 were mediated to a significant degree by a sodium ion current. The threshold for the Ins(1,4) P_2 response was > 0.4 pmol per oocyte and the injection of larger doses (2–18 pmol per oocyte) often led to an irreversible depolarization and deterioration of the cell (not shown).

Our results support the hypothesis that Ins(1,4,5) P_3 is the intracellular mediator of the muscarinic depolarizing chloride current. It is possible that Ins(1)P or Ins(1,4) P_2 are involved in the control of sodium ion gating, although nothing is known about neurotransmitter control of sodium channels in this cell. We cannot exclude the possibility that the effects of Ins(1)P and of Ins(1,4) P_2 injection are caused by impurities contained in these fractions. The very low threshold concentrations imply, however, that these are responses of considerable specificity.

This demonstration of a relatively simple assay system for the physiological effects of Ins(1,4,5) P_3 and its congeners offers a unique opportunity to extend the study of the effects of phosphoinositide breakdown products to the metabolism and molecular genetics of these compounds.

This research was supported by grants from the Schreiber Foundation and the Recanati Foundation for Medical Research. **Note added in proof:** It has been demonstrated recently^{17,18} that Ins(1,4,5)P₃ microinjection into *Limulus* photoreceptors mimics physiological responses in this organ.

Received 14 June; accepted 8 October 1984.

1. Berridge, M. J. *Biochem. J.* **212**, 849-858 (1983).
2. Berridge, M. J., Dawson, R. M. C., Downes, P. C., Heslop, J. P. & Irvine, R. F. *Biochem. J.* **212**, 473-482 (1983).
3. Streib, H., Irvine, R. F., Berridge, M. J. & Schulz, I. *Nature* **306**, 67-69 (1983).

Transmission distortion of *t*-haplotypes is due to interactions between meiotic partners

Anna W. Seitz & Dorothea Bennett

Memorial Sloan-Kettering Cancer Center, Sloan-Kettering Division, Graduate School of Medical Sciences, Cornell University, New York, New York 10021, USA

The *T/t*-complex of the mouse includes a series of recessive lethal and semi-lethal mutations but, despite such lethalties, mutant *t*-haplotypes are found in high frequency in wild mouse populations¹. This polymorphism is apparently maintained because heterozygous males preferentially transmit the *t*-bearing chromosome to their offspring². Despite many attempts to define the basis of the transmission ratio distortion³⁻¹⁰, it has been unclear whether this is because *t*-bearing sperm have better than average fertilizing ability or whether *+*-bearing sperm in heterozygous males are rendered defective. To examine this point, we constructed male (XY↔XY) chimaeras containing *+/+* and *+/t^{w73}* genotypes, marked respectively by albino and pigmented coat colours, and two isozyme variants. Such males produce a mixture of three different sperm types: *+*-bearing sperm from the *+/+* genotype, *+*-bearing sperm from the *+/t* genotype, and *t*-bearing sperm from the *+/t* genotype. Appropriate matings can distinguish between these three types, and our data, reported here, show that *t*-bearing sperm in chimaeric mice maintain their advantage over their 'meiotic partners' but do not have any advantage over sperm from the *+/+* genotype.

4. Agranoff, B. W., Murthy, P. & Seguin, E. B. *J. biol. Chem.* **258**, 2076-2078 (1983).
5. Putney, J. W. Jr, Burgess, G. M., Halenda, S. P., McKinney, J. S. & Rubin, R. P. *Biochem. J.* **212**, 483-488 (1983).
6. Poggioli, J., Weiss, S. J., McKinney, J. S. & Putney, J. W. Jr. *Molec. Pharmac.* **23**, 71-77 (1983).
7. Kusano, K., Miledi, R. & Stinnakre, J. *Nature* **270**, 739-741 (1977).
8. Dascal, N. & Landau, E. M. *Life Sci.* **27**, 1423-1428 (1980).
9. Dascal, N., Landau, E. M. & Lass, Y. *J. Physiol., Lond.* **352**, 551-574 (1984).
10. Dascal, N. & Landau, E. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3052-3056 (1982).
11. Robbins, N. & Molenaar, P. C. *Proc. R. Soc. B213*, 59-72 (1981).
12. Soreq, H., Parvari, R. & Silman, I. *Proc. natn. Acad. Sci. U.S.A.* **79**, 830-834 (1982).
13. Dascal, N., Yekuel, R. & Oron, Y. *J. exp. Zool.* **230**, 131-135 (1984).
14. Downes, P. C. & Michell, R. H. *Biochem. J.* **198**, 133-140 (1981).
15. Downes, P. C., Mussat, M. C. & Michell, R. H. *Biochem. J.* **203**, 169-177 (1982).
16. Bartlett, G. R. *J. biol. Chem.* **234**, 466-468 (1969).
17. Fein, A., Payne, R., Corson, D. W., Berridge, M. J. & Irvine, R. F. *Nature* **311**, 157-160 (1984).
18. Brown, J. E. *et al. Nature* **311**, 161-163 (1984).

Chimaeras were made by aggregating two eight-cell embryos; these were cultured and transferred at the morula or blastocyst stage to the uterus of a pseudopregnant female¹¹. They were allowed to develop to term and delivered naturally. The *+/+* embryos were (BALB/c×SJL)_{F1}, albino in coat colour (*c/c*) and homozygous for the isozymes phosphoglucose isomerase A (GPI-A) and glyoxylase 1A (GLO-1A). The *+/t^{w73}* genotype embryos were obtained from males with a transmission ratio >95%, and were pigmented (*c⁺/c⁺*) and homozygous for *Gpi-b*. The genetic marker for the *t*-haplotype was *Glo-1^a*, which is within its region of recombination suppression, whereas the *+* chromosome carried *Glo-1^b*. Thus, the *+/t* heterozygous genome in chimaeras could be identified by the presence of GLO-1A, GLO-1B and the hybrid AB band separable by electrophoresis.

Of 32 coat colour chimaeras obtained, seven males produced both albino and pigmented offspring in matings with *T/+* albino females. The pigmented progeny of three of these males included tailless (*T/t*), as well as short (*T/+*) and normal tailed (*t/+* or *+/+*) animals. The tailless progeny identify the presence of *t^{w73}* in the parental chimaera, while the short tailed progeny define them as also carrying the *+* allele; thus, the pigmented component of these chimaeras must be *t^{w73}/+*. The remaining four males produced pigmented progeny consisting only of short and normal tailed animals, and must therefore be *+/+*. Although there was considerable inter-male heterogeneity, both groups produced approximately equal proportions of pigmented and albino offspring (Table 1), thus suggesting that the two input genotypes of these chimaeras are equally efficient at sperm production, and that the presence or absence of *t^{w73}* in the

Table 1 Characterization and breeding behaviour of chimaeric *+/t^{w73}* ↔ *+/+* mice

Mouse	Relative proportion of coat colour pigmented:albino (<i>+/t</i> or <i>+/+</i>)(<i>+/+</i>)	Contribution of each genotype to testicular cells/sperm GPI-B:GPI-A (<i>+/t</i> or <i>+/+</i>)(<i>+/+</i>)	Offspring pigmented:albino	% Pigmented offspring	Tail phenotype of pigmented offspring nt:sh:ot	Transmission ratio of pigmented offspring (%)
<i>+/t</i> ↔ <i>+/+</i>						
4383	50:50	30:70	8:139	5	3:0:5	100
4387	50:50	70:30	163:80	67	72:2:89	98
4388	50:50	50:50	125:28	82	58:5:62	93
			Total 296:247	55		
<i>+/+</i> ↔ <i>+/+</i>						
424	90:10	ND	41:20	67	19:22:0	—
430	70:30	ND	37:9	77	23:19:0	—
4380	60:40	30:70	55:111	33	31:24:0	—
3973	50:50	ND	26:1	93	12:14:0	—
			Total 159:143	53		
			Grand total 455:390	54		

nt, Normal tail (*+/+*, *+/t*); sh, short tail (*T/+*); ot, tailless (*T/t*); ND, not determined. Genotypes (*+/+*, *+/t*) were determined by GLO-1 isozyme analysis and progeny testing. GLO-1 isozyme identification was by a method modified from ref. 18. Sperm from both epididymis and vas deferens were analysed for GPI isozymes. Epididymis and vas deferens were removed from freshly killed mice and cut into small pieces in sperm medium (150 mM NaCl, 10 mM sodium phosphate (pH 7.2, 0.3% fructose, 0.3% dextrose, 1 mM CaCl₂, 1 mM MgCl₂) to release the sperm. This was passed through 37-μm gauze and centrifuged at 400 r.p.m. for 5 min. The sperm in the suspension were pelleted, frozen, then thawed in a small amount of distilled water. The GPI isozyme ratio of this supernatant was determined by electrophoresis¹⁹. The GPI isozyme ratio of spermatogenic cells in the testis, prepared according to Silver *et al.*²⁰, was also determined and found to correspond to that of the sperm in the same chimaera.

chimaeric genotype is not a factor controlling the overall proportions of pigmented and albino sperm.

The transmission frequency of t^{w73} relative to its +homologue is measured by the proportion of tailless to short tailed offspring, and the transmission frequency of t^{w73} from chimaeric males remains very high (Table 1). Thus, t -bearing sperm maintain their superiority over their meiotic partner, although this superiority does not extend to the +bearing sperm from the +/+ genotype.

Our data therefore suggest that t -bearing sperm are not intrinsically superior, but that +bearing sperm produced by +/t heterozygotes are ineffective compared with both t -bearing sperm and +bearing sperm from another genotype. The defect is therefore in the +bearing sperm whose 'meiotic partners' were t -chromosomes. These results help to explain the apparent paradox that whereas heterozygous +/t and T/t males show distorted transmission ratios in favour of the t -bearing chromosome, heterozygous t^x/t^y males (where t^x and t^y are t -haplotypes belonging to different complementation groups) are sterile. This led to the hypothesis that t -bearing sperm somehow 'impair' sperm of the other haploid class, to result in a high transmission ratio in +/t or T/t heterozygotes and sterility in t^x/t^y heterozygotes⁴. However, when sperm from t^x/t^y males were mixed with those from +/+ or T/t males before artificial insemination, no deleterious effect on non- t -bearing sperm is observed⁴. Our demonstration of an impairing effect suggests that it must occur early in spermatogenesis. This agrees with previous observations that chimaeric males carrying +/t^{w2} ↔ t^{w5} genotypes and breeding for both genotypes transmit the characteristic high proportion of each haplotype relative to its '+' allele, with no loss of fertility⁵. Thus, both genotypes undergoing spermatogenesis in the same testis differentiate independently of each other.

Transmission distortion in the mouse seems to be similar, but not entirely analogous, to *Segregation Distortion* (SD) in *Drosophila*. Genetically the two systems are very similar, involving inversions, lethal genes and sterility, as well as 'responder' and 'distorter' loci¹²⁻¹⁵. However, the SD locus causes sperm carrying the homologous chromosome to degenerate during development in the testis¹², whereas in the mouse, +/t heterozygotes are known to produce both t -bearing and non- t -bearing sperm. Hammerberg and Klein⁶ found equal representation of t -bearing chromosomes and their homologues at meiosis I. Yanagisawa *et al.*⁷ found two serologically distinct populations of sperm in +/t males. Also, H.-S. Shin (personal communication) demonstrated by Southern blotting that equal proportions of t -bearing and +bearing sperm are found in both the epididymis and vas deferens. Furthermore, sperm from +/t heterozygotes show few morphological abnormalities^{16,17}. It is possible, therefore, that the presence of compromised +bearing sperm in sperm populations in heterozygous males may account for the differences revealed in other studies of viability and fertilization capability^{7,8}, differences in sperm mobility between t -carrying and wild-type mice⁹, and abnormally high levels of surface-associated galactosyltransferase¹⁰.

The deleterious action of the t -chromosome on its homologue seems likely to occur during the early stages of spermatogenesis, either during the diploid stage when the chromosomes are in the same nucleus, or by transfer of material through cytoplasmic bridges during spermiogenesis. The mechanism may operate at the DNA level during synapsis to alter gene expression on the homologous chromosome, or substances may be passed through cytoplasmic bridges between spermatocytes or spermatids to affect later sperm development and function of the non- t -bearing sperm.

We thank Katarina Eisinger for technical assistance. This work was supported by NIH grant HD 14526. A.W.S. is a recipient of a NIH National Research Service Award HD 06367.

Received 12 September; accepted 25 October 1984.

1. Bennett, D. *Cell* **6**, 411-454 (1975).
2. Dunn, L. C. *Am. Nat.* **94**, 385-393 (1960).
3. Bryson, V. J. *Morphol.* **74**, 131-187 (1944).
4. Braden, A. W. H. & Gluecksohn-Waelsch, S. *J. exp. Zool.* **138**, 431-452 (1958).

5. Papiannou, V. E., Artzt, K., Doohar, G. B. & Bennett, D. *Gamete Res.* **2**, 147-151 (1979).
6. Hammerberg, C. & Klein, J. *Nature* **253**, 137-138 (1975).
7. Yanagisawa, K., Pollard, D. R., Bennett, D., Dunn, L. C. & Boyse, E. A. *Immunogenetics* **1**, 91-96 (1974).
8. Braden, A. W. H. *Nature* **181**, 786-787 (1958).
9. Olds-Clarke, P. *Genet. Res.* **42**, 151-157 (1983).
10. Shun, B. D. & Hall, N. G. *J. Cell Biol.* **95**, 567-573 (1982).
11. Mintz, B. in *Methods in Mammalian Embryology* (ed. Daniels, J. C.) 186-214 (Freeman, San Francisco, 1971).
12. Peacock, W. J., Tokuyasu, K. T. & Hardy, R. W. *Proc. int. Symp. Genetics of the Spermatozoon* (eds Beatty, R. A. & Gluecksohn-Waelsch, S.) 247-268 (Bogtrykkeriet, Copenhagen, 1971).
13. Bennett, D., Alton, A. K. & Artzt, K. *Genet. Res.* **41**, 29-45 (1983).
14. Styra, M. J. & Klein, J. *Genet. Res.* **38**, 315-325 (1981).
15. Lyon, M. F. *Cell* **37**, 621-628 (1984).
16. Yanagisawa, K. *Japan. J. Genet.* **40**, 97-104 (1965).
17. Hillman, N. & Nadjicka, M. *J. Embryol. exp. Morph.* **44**, 243-261 (1978).
18. Harris, H. & Hopkinson, D. A. *Handbook of Enzyme Electrophoresis in Human Genetics* (North-Holland, Amsterdam, 1978).
19. Eicher, E. & Washburn, L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 946-956 (1978).
20. Silver, L. M., Artzt, K. & Bennett, D. *Cell* **17**, 275-284 (1979).

Amplification, enhanced expression and possible rearrangement of EGF receptor gene in primary human brain tumours of glial origin

Towia A. Libermann*, Harris R. Nusbaum*,
Nissim Razon†, Richard Kris*, Irit Lax*,
Hermona Soreq‡, Nigel Whittle§,
Michael D. Waterfield§¶, Axel Ullrich
& Joseph Schlessinger*¶

* Department of Chemical Immunology and ‡ Department of Neurobiology, The Weizmann Institute of Science, Rehovot 76100, Israel

† Department of Neurosurgery, Ichilov Hospital, Tel Aviv Medical Center, Tel Aviv 64239, Israel

§ Protein Chemistry Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

¶ Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

Epidermal growth factor (EGF), through interaction with specific cell surface receptors, generates a pleiotropic response that, by a poorly defined mechanism, can induce proliferation of target cells¹⁻⁴. Subversion of the EGF mitogenic signal through expression of a truncated receptor⁵⁻⁷ may be involved in transformation by the avian erythroblastosis virus (AEV) oncogene *v-erb-B*, suggesting that similar EGF receptor defects may be found in human neoplasias. Overexpression of EGF receptors has been reported on the epidermoid carcinoma cell line A431 (ref. 8), in various primary brain tumours⁹ and in squamous carcinomas¹⁰. In A431 cells the receptor gene is amplified^{6,11,12}. Here we show that 4 of 10 primary brain tumours of glial origin which express levels of EGF receptors that are higher than normal also have amplified EGF receptor genes. Amplified receptor genes were not detected in the other brain tumours examined. Further analysis of EGF receptor defects may show that such altered expression and amplification is a particular feature of certain human tumours.

Examination of eight novel brain tumours together with those 29 tumours studied previously, using a highly sensitive assay for EGF receptor kinase activity, suggested that some glioblastomas express high levels of kinase compared with other tumours and with normal brain material obtained postmortem⁹ (see Table 1). Because the A431 vulval carcinoma cell line which expresses high levels of EGF receptors⁸ also has amplified EGF receptor genes^{6,11-13}, we examined the structure of the receptor gene in DNA prepared from 21 brain tumours.

We have recently obtained cDNA clones from A431 cells and placenta; sequence analysis of these has provided the complete primary structure of the EGF receptor, together with the nucleotide sequence of the 5'- and 3'-untranslated regions of a 5.8-

¶ To whom correspondence should be addressed.

Table 1 Properties of human brain tumours

Tissue designation	Source of DNA	No. of EGF receptor gene copies	EGF receptor expression
Controls			
HPL	Human placenta	1	High
A431	Epidermoid vulval carcinoma	15-25	High
Brain tumours			
GM1	Glioblastoma multiforme	20-30	High
GM2(6)	Glioblastoma multiforme	10-20*/25-42†	High
GM3	Glioblastoma multiforme	1	ND
GM4	Glioblastoma multiforme	1	Low
GM5	Glioblastoma multiforme	1	ND
GM6	Glioblastoma multiforme	50-60	High
GM7 (20)	Glioblastoma multiforme	1	Low
GM8	Glioblastoma multiforme	1	None
GM9	Glioblastoma multiforme	6-9	ND
	Adjacent normal brain tissue	1	ND
GM10	Glioblastoma multiforme	1	ND
CGM1 (10)	Cerebellar glioblastoma	1	Low
CGM2	Cerebellar glioblastoma	1	ND
ACI11 (2)	Astrocytoma grade II	1	Medium
ACI12	Astrocytoma grade II	1	ND
ODG1	Oligodendroglioma	1	ND
ODG2	Oligodendroglioma	1	Low
ODG3	Oligodendroglioma	1	Medium
MEN1	Meningioma	1	Medium
MEN2	Meningioma	1	ND
MEN3	Meningioma	1	Medium
PNE	Primitive neuroectodermal tumour	1	ND

Tumour tissue was frozen in liquid nitrogen immediately after removal to avoid degenerative changes. As only small amounts could be obtained, the kinase assay provides the only practical assay for EGF receptor levels in these samples at present. The radioactive content of the ^{32}P -labelled EGF receptor was determined and expressed semi-quantitatively as follows: low, <2,000 c.p.m.; medium, 2,000-20,000 c.p.m.; high, >20,000 c.p.m. For further details see ref. 9. ND, not determined. Numbers in parentheses refer to tumours analysed previously⁹.

* 5' probe (II); † 3' probe (V) (see Figs 2, 3).

kilobase (kb) mRNA⁶. Various segments of the cDNA clones have been subcloned and used to generate probes which can detect restriction fragments of genomic DNA encoding defined regions of the 5.8-kb mRNA (Figs 1, 2). Southern blot hybridization analysis of equal amounts of *Eco*RI-digested DNA from the brain tumours, using a *Pst*-*Pst* fragment of the EGF receptor cDNA as a probe, showed that 4 of 10 glioblastomas contain amplified copies of the EGF receptor gene (Fig. 1, Table 1). Based on densitometric quantification of Southern blots, made with various amounts of DNA, we estimate that the EGF receptor gene is amplified 15-25-fold in A431 cells and 6-60-fold in the four glioblastomas compared with the gene in human placenta. The DNAs prepared from meningioma, oligodendroglioma or astrocytoma grade II tissue samples and normal brain tissue adjacent to tumour GM9 showed no amplification of the EGF receptor genes (Fig. 1, Table 1). To characterize the amplified receptor genes further, Southern blots of *Eco*RI- and *Hind*III-digested DNAs were probed sequentially with cDNA fragments corresponding to other regions of the 5.8-kb mRNA (see Fig. 2). Novel restriction fragments were detected in both the *Eco*RI and *Hind*III digests of DNA from tumour GM2 with probe III (1.3 kb with *Eco*RI and 9.4 kb with *Hind*III) and in the *Hind*III digest of tumour GM1 DNA with probe IV (3.4 and 5.3 kb) (Fig. 2). Furthermore, the relative intensities of restriction fragments detected by hybridization with distinct probes (I-IV) were significantly different between GM1, GM2 and A431 cell DNAs. For example, the relative intensities of the *Eco*RI restriction fragments detected in GM2, GM1 and A431 DNAs using probe IV were significantly different from those detected with probe III (see Fig. 2). The amplified genes in tumours GM6 and GM9 have not yet been analysed in detail.

These results are consistent with amplification and rearrangement of the receptor gene in these tumours, generating additional normal and aberrant genes, as reported for other amplified genes¹⁴. For the GM1 and GM2 genes, the detection of novel restriction fragments suggests that rearrangement involves the coding region; the nature of the rearrangement in A431 DNA

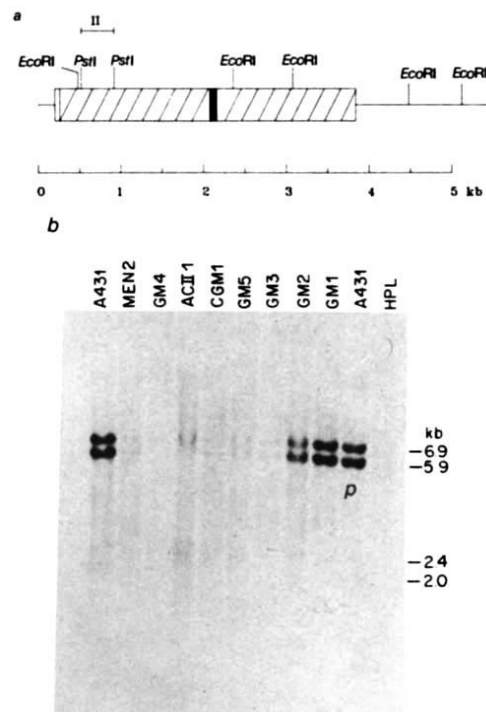


Fig. 1 Southern blot analysis of DNA from primary human brain tumours and human placenta. **a**, Schematic representation of human EGF receptor cDNA. Untranslated sequences are represented by a line; coding sequences are boxed. The open box represents sequences encoding the signal peptide; hatched regions encode the mature receptor. The black region demarcates the coding region for the putative transmembrane domain. The cDNA probes used for this study were constructed as described previously⁶, except for probe p8.4 (designated here as probe II) which is derived from the recombinant plasmid p8 (T.A.L. *et al.*, unpublished results). **b**, Amplification of the EGF receptor gene in glioblastoma and A431 epidermoid carcinoma cells and human placenta. The DNAs from these tumours were digested with *Eco*RI and probed with ^{32}P -labelled probe II DNA. GM1, GM2, GM3, GM4 and GM5 are different primary human glioblastomas, CGM1 is a cerebellar glioblastoma, ACI11 is an astrocytoma grade II, MEN2 is a meningioma, and A431 is a human epidermoid carcinoma cell line. HPL is human placenta which served as a control tissue. Some of these tumours were analysed previously for receptor kinase activity (see Table 1 and ref. 9).

Methods: **a**, Cloning of EGF receptor cDNA clone p8. mRNA from A431 cells was isolated by the guanidine thiocyanate/CsCl method²⁸. cDNA synthesis, cloning into pUC9 and colony screening were carried out according to published protocols²⁹. Briefly, a 17mer probe, as a mixture of 256 oligonucleotides (3'ATA/GTTA/G GGX TGX TGX AT 5') based on the amino acid sequence of a tryptic peptide of A431 EGF receptor (ref. 5) was end-labelled with T4 polynucleotide kinase (New England Biolabs) and (γ - ^{32}P)ATP (3,000 Ci nmol⁻¹; Amersham). Colony screening for hybridization was carried out in 6×SSC, 5×Denhardt's, 0.1% SDS, 100 $\mu\text{g ml}^{-1}$ salmon sperm DNA at room temperature for 36 h. Filters were washed at 45 °C with 3×SSC, 0.1% SDS. Nucleotide sequence analysis was carried out according to the procedure of Maxam and Gilbert³⁰. p8 is a 2.5-kb cDNA clone derived from the 2.8 kb variant mRNA from A431 cells described by Ullrich *et al.*⁶. Probe II is a *Pst* fragment (399 base pairs, (bp)) derived from clone p8. **b**, DNA of high relative molecular mass was isolated as described previously²⁸, and 15 μg were digested to completion with excess *Eco*RI or *Hind*III (New England Biolabs), fractionated by electrophoresis through a 0.7% agarose gel and transferred to nitrocellulose paper. Probe II was radiolabelled with [α - ^{32}P]dATP and [α - ^{32}P]dCTP (Amersham) by the procedure of Taylor *et al.*³¹. Hybridization with 10⁷ c.p.m. of ^{32}P -labelled probe was performed in 6×SSC, 5×Denhardt's, 10% dextran sulphate, 50 mM sodium phosphate pH 6.5 and 100 $\mu\text{g ml}^{-1}$ salmon sperm DNA at 65 °C for 16 h. Filters were washed in 0.2×SSC, 0.1% SDS at 65 °C and autoradiographed for 1 day at -70 °C using intensifier screens. Sizes were calculated using λ phage DNA cleaved with restriction endonuclease *Hind*III as standards.

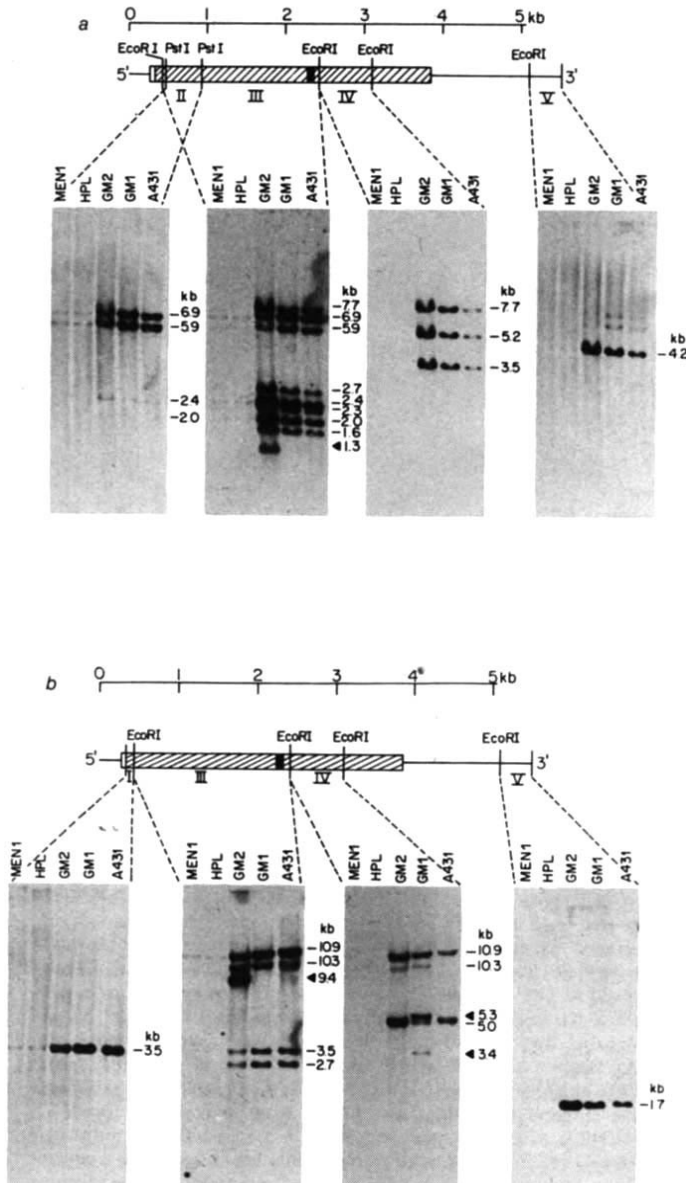
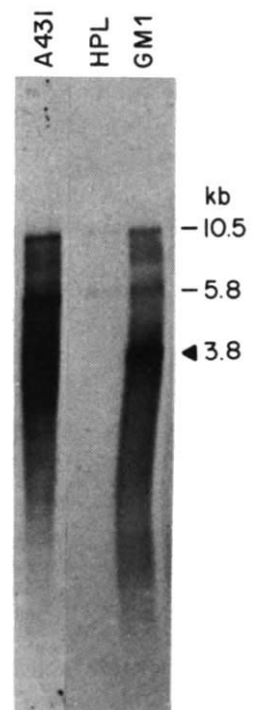


Fig. 2 Southern blot analysis of EGF receptor sequences in human brain tumours, A431 cells and human placenta. A schematic representation of the receptor mRNA and its coding sequences is shown. Dashed lines indicate the regions of the schematically depicted cDNA giving rise to the hybridization patterns shown (see also Fig. 1). Arrowheads indicate the location of fragments of DNA in the amplified EGF receptor gene of the glioblastoma tumours, that were not detected in other DNAs. Roman numerals (I-V) denote cDNA fragments used as hybridization probes (see Fig. 1). High-molecular weight DNA (15 µg) from A431 cells, human placenta (HPL), glioblastomas (GM1, GM2) and meningioma (MEN1) was digested with *EcoRI* (a) or *HindIII* (b). To exclude technical problems due to incomplete digestion, the digestions were repeated several times with a large excess of restriction enzymes, revealing the same results. The DNAs were electrophoresed and blotted as described in Fig. 1 legend. The blots were hybridized to the nick-translated EGF receptor-specific cDNA inserts isolated from relevant cDNA clones under high stringency conditions, as described in Fig. 1 legend. The same blots were re-used for the different probes without detectable loss of signal. To denature the cDNA-genomic DNA hybrids for re-use, filters were soaked with slight agitation in 0.5 M NaOH, 1.5 M NaCl for 10 min at room temperature, rinsed with water, neutralized twice for 10 min in 0.5 M Tris-HCl pH 7.0, 1.5 M NaCl and washed in 3×SSC. Filters were kept wet in Saran wrap, preincubated in hybridization buffer and re-used as described above.

Fig. 3 Northern blot analysis of mRNA from A431 cells, human placenta and glioblastoma. A431 cells were grown in Dulbecco's modified Eagle's medium containing 10% fetal calf serum in an atmosphere of 5% CO₂/95% air at 37 °C. RNAs were isolated from frozen tissue of human placenta, from glioblastoma GM1 after pulverization in liquid N₂ and from fresh A431 cells by the guanidine thiocyanate/CsCl method described elsewhere²⁸. Aliquots (10 µg) of poly(A)-selected mRNA were heated at 60 °C for 10 min in a solution containing 50% formamide, 6% formaldehyde and running buffer (20 mM MOPS pH 7.0, 5 mM Na-Ac, 1 mM EDTA). The samples were electrophoresed at 100 V for 4 h in 1% agarose gels containing 6% formaldehyde and 1×running buffer. The RNA was transferred with 10×SSC to nitrocellulose filters, fixed by heating at 80 °C for 2 h and hybridized at 42 °C for 2 days with 2×10⁶ c.p.m. ml⁻¹ of nick-translated p64.3 probe in a solution containing 50% formamide, 5×SSC, 10% dextran sulphate, 1×Denhardt's mixture, 10 mM sodium phosphate pH 6.8 and 100 µg ml⁻¹ salmon sperm DNA. After washing at 65 °C with 0.1×SSC, 0.1% SDS, the filters were autoradiographed for 2 days at -70 °C using intensifier screens.



is less clear. The novel fragments could have arisen from polymorphic genes, but this is unlikely as fragments of similar size were not seen in the other tumour DNAs.

Owing to the limited amounts of tissue available, RNA could be prepared only from tumour GM1. Northern blot hybridization analysis of GM1 and A431 RNAs with probe IV revealed the 5.8- and 10.5-kb mRNAs detected previously in placenta^{6,11,12}, together with several abnormal transcripts, notably a 3.8-kb species present at about 10 times the level of the normal transcripts in tumour GM1 (Fig. 3). Analysis of RNA with probe III also detected the 3.8-kb RNA in tumour GM1 and the 2.8-kb transcript which encodes a truncated external receptor domain present at ~100 times the level of the normal transcripts as characterized previously in A431 cells^{6,15,16} (data not shown). The abnormal transcripts may be encoded by the rearranged amplified genes. Further analysis is needed to determine the relationships between gene structure, aberrant mRNA expression and receptor expression levels in tumours, and problems inherent in analysis of small tissue samples must be overcome.

Amplifications of sequences related to oncogenes in primary human tumours has been observed for *c-myc*¹⁷, *N-myc*¹⁸⁻²⁰ and *Ha-ras1* (ref. 21), and amplification of sequences related to *abl* has been reported in a human cell line²². Karyotypic abnormalities including translocations and in some cases increased copies of chromosome 7, to which the EGF gene maps, have been documented in A431 cells^{23,24} and in several glioblastomas²⁵⁻²⁷ and could be related to amplification and/or rearrangement of the receptor gene.

Our results, although preliminary, raise several important questions regarding the expression of the EGF receptor in specific tumours. Because the *v-erb-B* oncogene may transform cells through expression of a truncated receptor^{5,6}, our observation of amplification and rearrangement of the receptor gene, together with abnormal transcription, in certain human tumours warrants further study as abnormal receptor expression could be involved in neoplasia.

We thank Dr Boaz Amit for the preparation of the oligonucleotide probe and Dr A. D. Bartal and his staff in the Department of Neurosurgery, Ichilov Hospital, Tel Aviv for assistance in obtaining and identifying the brain tumours. This work was

supported by grants from the NIH (CA-25820 to J.S.), the US-Israel Binational Science Foundation (to J.S.), and the Hermann and Lilly Schilling Foundation for Medical Research (to H.S.). We acknowledge discussions with Dr Moshe Oren, Dr Benny Shilo, Dr Cynthia Webb, David Wolff and Dr Mia Horowitz.

Received 1 August; accepted 26 October 1984.

- Cohen, S. J. *biol. Chem.* **237**, 1555-1562 (1962).
- Schlessinger, J. et al. *Membranes in Growth and Development* 359-372 (Liss, New York, 1982).
- Carpenter, G. & Cohen, S. A. *Rev. Biochem.* **48**, 193-216 (1979).
- Schlessinger, J. et al. *CRC crit. Rev. Biochem.* **14**, 93-111 (1983).
- Downward, J. et al. *Nature* **307**, 521-527 (1984).
- Ullrich, A. et al. *Nature* **309**, 418-425 (1984).
- Yamamoto, et al. *Cell* **35**, 71-78 (1983).
- Fabricant, R. N., DeLarco, J. E. & Todaro, G. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 565-569 (1977).
- Libermann, T. A. et al. *Cancer Res.* **44**, 753-760 (1984).
- Cowley, G., Smith, J. A., Gusterson, B., Hendler, F. & Ozanne, B. in *Cancer Cells I* (eds Levine, A. J. et al.) 5-10 (Cold Spring Harbor Laboratory, New York, 1984).
- Merlino, G. T. et al. *Science* **224**, 417-419 (1984).
- Lin, C. R. et al. *Science* **224**, 843-848 (1984).
- Xu, Y.-H. et al. *Nature* **309**, 806-810 (1984).
- Schimke, R. J. *Cell* **37**, 705-715 (1984).
- Mayes, E. L. V. & Waterfield, M. D. *EMBO J.* **3**, 531-537 (1984).
- Weber, W., Gill, G. N. & Spiess, J. *Science* **224**, 294-297 (1984).
- Dalla-Favera, R., Wong-Staal, F. & Gallo, R. C. *Nature* **299**, 61-63 (1982).
- Kohl, N. E. et al. *Cell* **35**, 359-367 (1983).
- Schwab, M. et al. *Nature* **305**, 245-248 (1983).
- Schwab, M. K. et al. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4940-4944 (1984).
- Hayashi, K., Kakizoe, T. & Sugimura, T. *Gann* **74**, 798-901 (1983).
- Collins, S. & Groudine, M. *Nature* **298**, 679-681 (1982).
- Shimizu, N., Behzadian, M. A. & Shimizu, Y. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3600 (1980).
- Davies, R. L., Grosse, V. A., Kucherlapati, R. & Bothwell, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4188-4192 (1980).
- Shapiro, J. R., Yung, W.-K. A. & Shapiro, W. R. *Cancer Res.* **41**, 2349-2359 (1981).
- Bartal, A. D., Heilbronn, Y. D. & Schiffer, J. *Surg. Neurol.* **1**, 337-342 (1973).
- Mark, J., Westermarck, B., Ponten, J. & Hugosson, R. *Hereditas* **87**, 243-260 (1977).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).
- Helfman, D. M., Feramisco, J. R., Fiddes, J. C., Thomas, G. P. & Hughes, S. H. *Proc. natn. Acad. Sci. U.S.A.* **8**, 31-35 (1983).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
- Taylor, J. M., Illmensee, R. & Summers, J. *Biochim. biophys. Acta* **4**, 324-330 (1976).

A hypersensitive site in *hsp70* chromatin requires adjacent not internal DNA sequence

Nancy A. Costlow, Jeffrey A. Simon & John T. Lis

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Wing Hall, Ithaca, New York 14853, USA

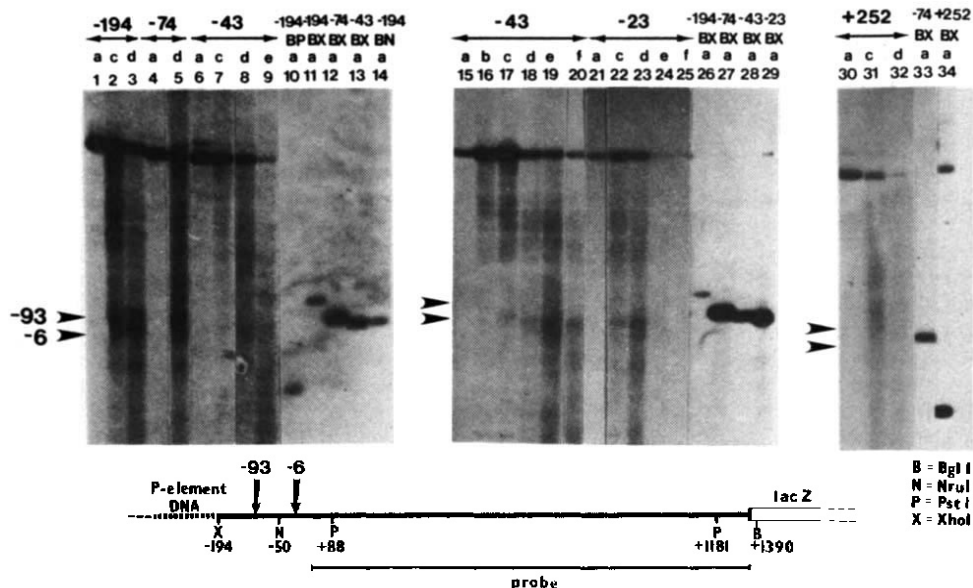
Nuclease-hypersensitive sites in chromatin exist at the 5' side of many eukaryotic genes. To gain some understanding of the molecular basis of these hypersensitive sites, we have now examined the pair of sites upstream of the *Drosophila hsp70* gene in a series of plasmids that contain deletions in the hypersensitive region and have been transformed into yeast cells. Hypersensitive sites 5' to a *Drosophila hsp70* gene are preserved when this gene is introduced into yeast by transformation¹. We find that a yeast strain containing a plasmid in which the deletion extends through the first

hypersensitive site still displays the normal pair of hypersensitive sites, so DNA sequences over which the first hypersensitive site is centred are not required for hypersensitivity at this position and the site can form over a foreign DNA sequence juxtaposed against this deletion end point. Deletions progressing further into the region bracketed by the pair of 5' hypersensitive sites eliminate the first hypersensitive site and alter the downstream site. We propose that the hypersensitive sites are generated through the binding of a protein that renders flanking sequences more accessible to nucleases, perhaps by preventing normal chromatin packaging.

When either the *hsp70* or *hsp83* gene of *Drosophila* is integrated into the yeast genome, the pattern of hypersensitivity 5' to these genes remains identical to that in *Drosophila*¹. The hypersensitive sites also occur on copies of the *hsp70* gene inserted into an episomal, centromere-containing yeast plasmid (Fig. 1). This plasmid, SP1.1, contains one copy of an *hsp70-lacZ* fusion gene² in which the *hsp70* sequences extend upstream to

Fig. 1 Detection of DNase I hypersensitive sites 5' to deletion mutants of the *Drosophila hsp70* gene transformed into yeast. The SP1.1 plasmid is a yeast centromere-containing vector that possesses one copy of an *hsp70-lacZ* fusion gene² flanked by P-element DNA⁴. The plasmid is a derivative of pYE(CEN3)41 (ref. 11) and contains the *hsp70-lacZ* fusion gene as in ep19.1 (ref. 2). The restriction map shows the relevant regions of the plasmid SP1.1. Numbering is from the *Drosophila* start of transcription of the *hsp70* gene. Heavy line, *hsp70* DNA; open line, *lacZ* sequences; broken line, P-element DNA. Deletions were made from the *XhoI* site with *Bal31* nuclease¹², *XhoI* linkers were added, and except for the -23 bp deletion in which about 7 bp of P-element DNA was also deleted, the plasmid was reconstructed so that the downstream deletion termini were rejoined to the *XhoI* site as in the parent plasmid.

The deletions studied are shown schematically in Fig. 3. The deletion termini for the -145 and +252 deletions are known ± 10 bp from high resolution analysis on agarose and acrylamide gels, but have not been sequenced. The -74, -43 and -23 deletion end points are known precisely from DNA sequence analysis (B. Wills and P. Wensink, personal communication). DNA was isolated from nuclei of yeast transformed with these plasmids, restriction cut with *BglI* and *EcoRI*, electrophoresed on a 1.4% agarose gel and hybridized to a ³²P-labelled *Bam*HI, *SalI* DNA fragment homologous to the 5' half of the *Drosophila hsp70* gene as previously described¹ and as depicted in the figure. The numbers above each lane indicate the deletion examined. The preparation of the -43 deletion used in lanes 15-20 was independent of that used in lanes 6-9, 13 and 28. DNA was isolated from yeast nuclei treated with DNase I at concentrations of 0 U ml⁻¹ (a), 10 U ml⁻¹ (b), 35 U ml⁻¹ (c), 100 U ml⁻¹ (d), 310 U ml⁻¹ (e) and 510 U ml⁻¹ (f). Arrowheads mark the positions of fragments resulting from cleavages at -93 and -6 bp. Additional restriction digests of non-DNase-treated samples were performed for molecular weight markers and are denoted with single-letter abbreviations for the restriction enzymes above the lanes: B, *BglI*; N, *NruI*; P, *PstI*; and X, *XhoI*. The intact *BglI* fragment, which is 3 kb in SP1.1, also hybridizes to the labelled DNA.



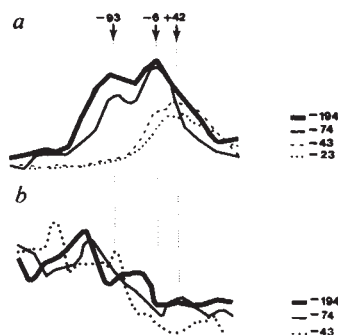


Fig. 2 Densitometry scans of hypersensitive regions 5' to *hsp70* in SP1.1 and its deletions. *a*, Densitometry scans of autoradiograms showing the hybridization of DNA isolated from DNase I-treated nuclei to an *hsp70*-specific restriction fragment: heavy line, the -194 bp deletion (the SP1.1 plasmid); light line, the -74 bp deletion; dashed line, the -43 bp deletion; dotted line, the -23 bp deletion. *b*, Densitometry scans of autoradiograms showing the hybridization of DNase I-treated purified plasmid DNA with an *hsp70*-specific restriction fragment. DNA was linearized by digestion with *Bgl*I and then treated with DNase I. The pattern obtained with the -23 deletion (not shown) was very similar to that of the -43 deletion. The arrows and the vertical dotted lines mark the positions of the DNase I hypersensitive sites found 5' to the *Drosophila* gene integrated into the yeast genome¹ at -93, -6 and +42 bp.

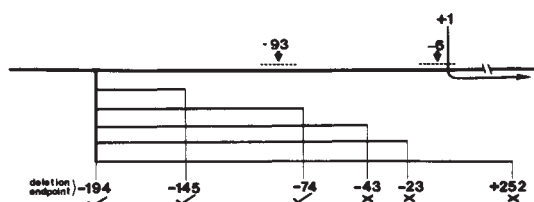


Fig. 3 Summary of *hsp70* deletion analysis in yeast. The two DNase I hypersensitive sites 5' to the *Drosophila hsp70* gene in the plasmid SP1.1 (the -194 deletion) are shown centred at -93 and -6 bp (± 10 bp). The start of transcription is designated as +1. Checks denote strains in which the two parental DNase I sites were observed; 'x', strains in which the parental hypersensitivity was not observed.

the *Xho*I site at -194 base pairs (bp). A series of mutant plasmids containing deletions extending from the *Xho*I site towards the gene were also examined for DNase I hypersensitive sites. The results show that deletion of the sequences between -194 and either -145 (data not shown) or -74 bp (Fig. 1, lane 5) does not affect the 5' hypersensitivity. Densitometry scans were taken of the hypersensitive regions 5' to *hsp70* in SP1.1 (Fig. 2a). The results with the deletion to -74 bp indicate that the sequence over which the DNase I site normally forms does not itself specify the hypersensitivity of this site.

Deletions to -43, -23 or +252 bp from the *Drosophila* start of transcription result in the abolition of the parental pattern of two 5' DNase I hypersensitive sites. In the +252 bp deletion, no hypersensitivity is observed over the region 5' to the coding sequences (Fig. 1, lanes 31, 32). In the -43 and -23 bp deletions, the two sites at -93 and -6 bp have been replaced by a broader region of hypersensitivity centred at +42 bp (Fig. 1, lanes 8, 17-20, 22-23). Superimposed densitometric scans of these autoradiograms (Fig. 2a) show clearly that the -93 site is absent and that the -6 site is reduced, although we cannot rule out some cutting at this site because of the overlap with the broad +42 peak. The +42 site may correspond to a minor site previously seen in *Drosophila* cells^{1,3}. Figure 3 summarizes these results.

As controls for sequence specificity of DNase I, purified plasmid DNA of each deletion was treated with DNase I over a range of DNase I concentrations and analysed in parallel with DNA from DNase I-treated nuclei. The results (Fig. 2) show

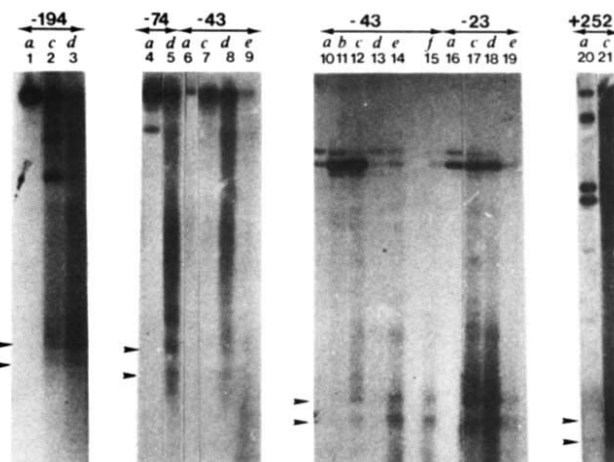


Fig. 4 Detection of DNase I hypersensitive sites 5' to the yeast *his4* gene in the strains carrying the *hsp70* deletions. The nitrocellulose filter shown in Fig. 1 was washed at 80 °C in 0.2 $\times$ SSC, 0.1% SDS, and hybridized with a fragment purified from the pFCI plasmid homologous to the yeast *his4* gene (F. Chumley and G. Fink, personal communication). DNA was isolated from yeast nuclei treated with DNase I at 0 U ml⁻¹ (*a*), 10 U ml⁻¹ (*b*), 35 U ml⁻¹ (*c*), 100 U ml⁻¹ (*d*), 310 U ml⁻¹ (*e*) or 510 U ml⁻¹ (*f*). Arrowheads mark the positions of fragments resulting from cleavages at the hypersensitive sites 5' to the *his4* gene. DNA in lanes 1-3 was isolated from the -194-bp deletion; lanes 4 and 5, the -74-bp deletion; lanes 6-9 and 10-15, two different preparations of the strains carrying the -43 bp deletion; lanes 16-19, the -23 bp deletion; and lanes 20 and 21, the +252 bp deletion. These are the same DNA samples as are shown in Fig. 1.

that the DNase I hypersensitive sites detected in the digests of nuclei at positions -93, -6 and +42 bp do not result from sequence specificity of DNase I, but rather seem to be a consequence of chromatin structure. As an additional control, hybridizations were performed to show that the known hypersensitive sites 5' to the yeast *his4* gene could be detected in the DNA samples from intact nuclei which had been used to examine hypersensitive sites in the deletion mutants (Fig. 4).

In these plasmid constructions, P-element DNA⁴ replaces the *Drosophila* sequences normally found upstream of the *hsp70* gene. It is unlikely that this P-element DNA specifies the hypersensitive sites observed in the -145 and -74 bp deletions, as the -43, -23 and +252 bp deletions lack these sites. Examination of the junction sequences formed by P-element DNA, the *Xho*I linkers used and the *Drosophila hsp70* DNA at the deletion end points show that neither the original *hsp70* sequence nor a similar sequence was recreated in strains carrying any of these deletions.

Studies of the DNA sequences affecting the SV40-hypersensitive sites⁵ reveal the importance of the 21-bp repeated region which, like the region we have identified, is itself resistant to nuclease and resides between two hypersensitive regions. The mechanism by which these sequences specify hypersensitivity is not known. Involvement of non-B-form DNA such as Z-DNA or cruciforms has been hypothesized⁶. Alternatively, the creation of DNase I hypersensitive sites may be a consequence of interactions between the DNA sequence and a protein factor⁷. Wu³ has identified a protease-sensitive factor in isolated nuclei from non-heat-shock treated cells which blocks the progress of *Exo*III digestion in the region between hypersensitive sites (-40 to -12 bp). In addition, transcription factors isolated from *Drosophila* cells have been reported to bind specifically to sequences in and between the *hsp70* hypersensitive sites⁸. Pelham⁹ has observed that a sequence centred at -55 bp from the start of transcription of the *hsp70* gene, and found upstream from other *Drosophila* heat shock genes, is required for heat shock induction of the *Drosophila hsp70* gene in monkey COS

cells. The lack of DNase I hypersensitive sites in the -43 and -23 bp deletions could result from deletion of a necessary protein-binding site. The *Drosophila hsp70* gene is transcribed in yeast¹⁰ and appears to be induced approximately fourfold in response to heat shock (J. de Banzie and J.T.L., unpublished). Assuming that regulation of this gene in yeast is qualitatively similar to that in other heterologous systems, our results agree with the hypothesis that DNase I hypersensitive sites are generated through a sequence element that is required for regulated expression. Resolution of the precise relationship between chromatin structure and gene expression of this *Drosophila* gene will require a thorough analysis of similar deletions stably transformed in *Drosophila*.

We thank V. Vogt and E. Keller for valuable discussions and for comments on the manuscript. This work was supported by NIH grant GM25232.

Received 23 July; accepted 16 October 1984.

1. Costlow, N. A. & Lis, J. T. *Molec. cell. Biol.* **4**, 1853-1863 (1984).
2. Lis, J. T., Simon, J. A. & Sutton, C. A. *Cell* **35**, 403-411 (1983).
3. Wu, C. *Nature* **309**, 229-234 (1984).
4. O'Hare, K. & Rubin, G. M. *Cell* **34**, 25-35 (1983).
5. Jongstra, J. *et al. Nature* **307**, 708-714 (1984).
6. Groudine, M. & Weintraub, H. *Cell* **30**, 131-139 (1984).
7. Emerson, B. M. & Felsenfeld, G. *Proc. natn. Acad. Sci. U.S.A.* **81**, 95-99 (1984).
8. Parker, C. & Topol, J. *Cell* **37**, 273-283 (1984).
9. Pelham, H. R. B. *Cell* **30**, 517-528 (1982).
10. Lis, J. *et al. Heat Shock from Bacteria to Man* (eds Schlesinger, M. J., Ashburner, M. & Tissières, A.) 57-62 (Cold Spring Harbor Laboratory, New York, 1982).
11. Clarke, L. & Carbon, J. *Nature* **287**, 504-509 (1980).
12. Gray, H. B., Ostrander, D. A., Hodnett, J. L., Legerski, R. J. & Roberson, D. L. *Nucleic Acids Res.* **2**, 1459-1492 (1975).

Synthesis in *E. coli* of α_1 -antitrypsin variants of therapeutic potential for emphysema and thrombosis

Michael Courtney, Sophie Jallat, Luc-Henri Tessier, Annie Benavente, Ronald G. Crystal* & Jean-Pierre Lecocq

Transgene SA, 11, Rue de Molsheim, 67000 Strasbourg, France

* Pulmonary Branch, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland 20205, USA

The primary function of α_1 -antitrypsin (α_1 -AT), an antiprotease produced by the liver, is the inhibition of neutrophil elastase, a protease capable of hydrolysing most connective tissue components¹⁻⁴. The importance of α_1 -AT is demonstrated by the high incidence of early-onset emphysema in individuals with hereditary α_1 -AT deficiency (Type PiZZ), in whom serum levels of α_1 -AT are 10-20% of normal⁵⁻¹⁰. Oxidants in tobacco smoke can inactivate α_1 -AT *in vitro*^{11,12}, and studies have shown that α_1 -AT from the lungs of individuals who smoke cigarettes may also be partially inactivated, perhaps explaining the high incidence of emphysema associated with cigarette smoking^{13,14}. Oxidative inactivation is probably due to modification of the Met residue (Met³⁵⁸) at the P1 subsite position of the elastase binding site of the protein^{15,16}. To study the possibility of modulating the biological properties of α_1 -AT, we have introduced selected sequence modifications at the reactive site by *in vitro* mutation of a cloned α_1 -AT complementary DNA. We describe here the characterization of two α_1 -AT analogues produced in *Escherichia coli*. The first, α_1 -AT(Met³⁵⁸→Val), is not only fully active as an elastase inhibitor but is also resistant to oxidative inactivation. The other, α_1 -AT(Met³⁵⁸→Arg), no longer inhibits elastase but is an efficient thrombin inhibitor. The active site of the latter is identical to that of the α_1 -AT (Pittsburgh) variant, which was associated with a fatal bleeding disorder^{17,18}.

Studies on the reactivity of various serine proteases towards synthetic oligopeptide substrates have identified subsite

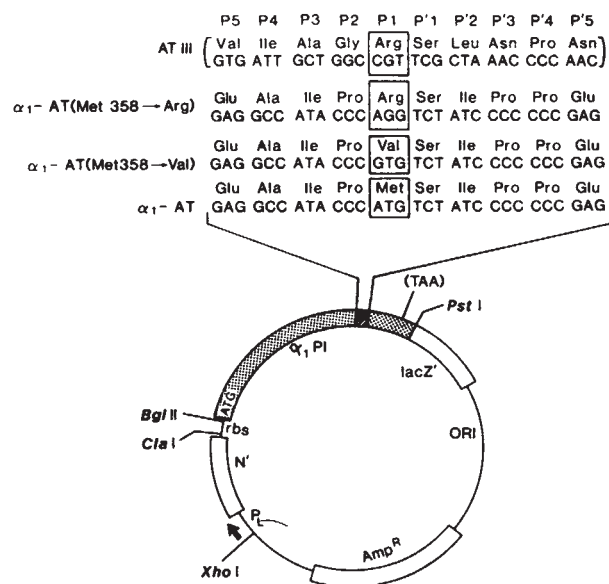


Fig. 1 Structure of the *E. coli* expression vector and active site sequences of α_1 -AT, antithrombin III, α_1 -AT(Met³⁵⁸→Val) and α_1 -AT(Met³⁵⁸→Arg). Transcription from the promoter P_L is controlled by the host-encoded temperature-sensitive repressor *cI*857 (ref. 20). Translation is initiated at a chemically synthesized ribosome binding site which was inserted downstream from P_L at the *Hpa*I site in the *N* gene³¹. The host used was *E. coli* TGE900 (*F*⁻*su*⁻*ilv*⁻*his*⁻*bio*(Δ *cI*857 Δ Bam Δ H1)). The dotted area represents mature α_1 -AT protein coding sequences and the inner cross-hatched zone corresponds to the α_1 -AT reactive site. The β -lactamase gene, the truncated *N* gene (*N'*) and the *lacZ'* gene are represented by the open areas. Standard nomenclature is used for the subsite positions.

Methods: α_1 -AT variants were obtained by oligonucleotide-directed mutagenesis essentially by methods already described^{22,33}. Briefly, the α_1 -AT-coding fragment *Bgl*II-*Pst*I was subcloned into M13mp701 (D. R. Bentley, personal communication) and single-strand phage DNA was hybridized to the oligonucleotides described in the text. Oligonucleotide-primed second-strand synthesis was carried out using DNA polymerase I (Klenow fragment) (Boehringer Mannheim) and the double-strand product then used to transform *E. coli* JM103. Mutant-containing plaques were identified by screening in stringent conditions (3×SSC, 50°C) with the original oligonucleotides. The nucleotide sequence of the mutant clones was verified using the dideoxynucleotide chain termination method²⁴. General methods and the use of the various enzymes have been described elsewhere^{20,31,35}.

sequence preferences for the various proteases¹⁹. As α_1 -AT and other inhibitors act as substrate analogues^{1,2}, such studies suggest that minor sequence changes at the inhibitor active site should drastically alter its reactivity or specificity. These studies¹⁹ showed that both human neutrophil elastase and porcine pancreatic elastase prefer a Val to a Met P1 residue at the substrate cleavage site. Because, unlike Met, Val cannot be oxidized, substitution of Val for Met at the α_1 -AT reactive site should give a protein that has the same elastase inhibitory capacity while being resistant to oxidative inactivation. To test this hypothesis, we have produced α_1 -AT(Met³⁵⁸→Val) in *E. coli*. We have also recreated the naturally occurring α_1 -AT (Pittsburgh) mutation, (Met³⁵⁸→Arg), and produced this modified protein in *E. coli*.

The isolation of a human α_1 -AT cDNA clone and its expression in *E. coli* have been described elsewhere²⁰. Efficient production of the α_1 -AT variants in *E. coli* was achieved by using the expression vector shown in Fig. 1. The unmodified α_1 -AT gene fragment was removed from the expression plasmid and subcloned into a M13 phage vector. Single-stranded phage DNA was used as target for oligonucleotide-directed mutagenesis of the α_1 -AT sequence. The oligonucleotides used, 5'-ATAGACACGGGTATG-3' and 5'-ATAGACCTGGGTATG-3',

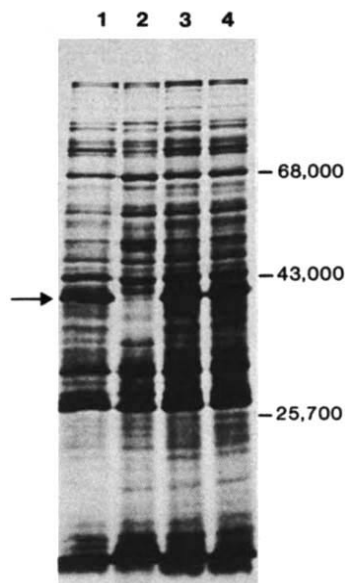


Fig. 2 Polyacrylamide gel electrophoresis of labelled extracts of *E. coli* producing modified and unmodified α_1 -AT. α_1 -AT synthesis was achieved by shifting the temperature of cultures (at $\sim 10^8$ cells ml^{-1}) from 28 to 37 °C. Aliquots of cultures, induced at 37 °C for 5 h, were labelled for 1 min with ^{35}S -methionine (Amersham). Samples of total cell extracts were analysed on 10% SDS-polyacrylamide gels followed by fluorography and autoradiography. Lane 1, extract containing unmodified α_1 -AT; lane 2, control extract (expression vector alone); lane 3, extract containing α_1 -AT(Met 358 $\rightarrow$ Val); lane 4, extract containing α_1 -AT(Met 358 $\rightarrow$ Arg). The arrow marks the position of α_1 -AT.

were complementary to the region encoding the α_1 -AT active site and each contained the appropriate mismatch at codon 358. Both variants involved substitution of one nucleotide only: ATG $\rightarrow$ GTG for Met 358 $\rightarrow$ Val and ATG $\rightarrow$ AGG for Met 358 $\rightarrow$ Arg (Fig. 1). After verification of the nucleotide sequence, the mutated genes were transferred back to the expression plasmid and examined for expression.

Polyacrylamide gel electrophoresis of *E. coli* extracts pulse-labelled with ^{35}S -methionine showed that α_1 -AT(Met 358 $\rightarrow$ Val) and α_1 -AT(Met 358 $\rightarrow$ Arg) were synthesized efficiently and at the same level as unmodified α_1 -AT (Fig. 2). In each case a strong band of equal intensity was induced which corresponded in size to nonglycosylated mature α_1 -AT (relative molecular mass 42,000) and which was immunoprecipitated with antiserum raised against plasma α_1 -AT (not shown). Radial immune diffusion assay (not shown) of sonicated culture supernatants demonstrated that the modified and unmodified α_1 -AT proteins accumulated to comparable levels ($\sim 3 \times 10^5$ molecules per cell).

Cleared sonicated extracts prepared from induced cultures containing either unmodified α_1 -AT or α_1 -AT(Met 358 $\rightarrow$ Val) were tested for their ability to inhibit human neutrophil elastase and porcine pancreatic elastase. The inhibition curves (Fig. 3a,b) indicate that α_1 -AT(Met 358 $\rightarrow$ Val) retained the ability to inactivate both elastases efficiently. Although plasma α_1 -AT, oxidized *in vitro*, is no longer able to inhibit porcine pancreatic elastase, there is significant residual inhibition of human neutrophil elastase ($k_{\text{ass}} = 3 \times 10^4 \text{ m}^{-1} \text{ s}^{-1}$, compared with $6.5 \times 10^7 \text{ m}^{-1} \text{ s}^{-1}$ for the unoxidized form)¹⁶. For this reason, the effect of oxidation of the *E. coli*-produced inhibitors was tested with the porcine enzyme. Oxidation of unmodified α_1 -AT using *N*-chlorosuccinimide led to complete inactivation of the protein (Fig. 3c), whereas in the same conditions, α_1 -AT(Met 358 $\rightarrow$ Val) retained almost full antielastase capacity (Fig. 3d). This confirms that oxidative inactivation is caused primarily by modification of the reactive site Met. It appears, therefore, that α_1 -AT(Met 358 $\rightarrow$ Val)

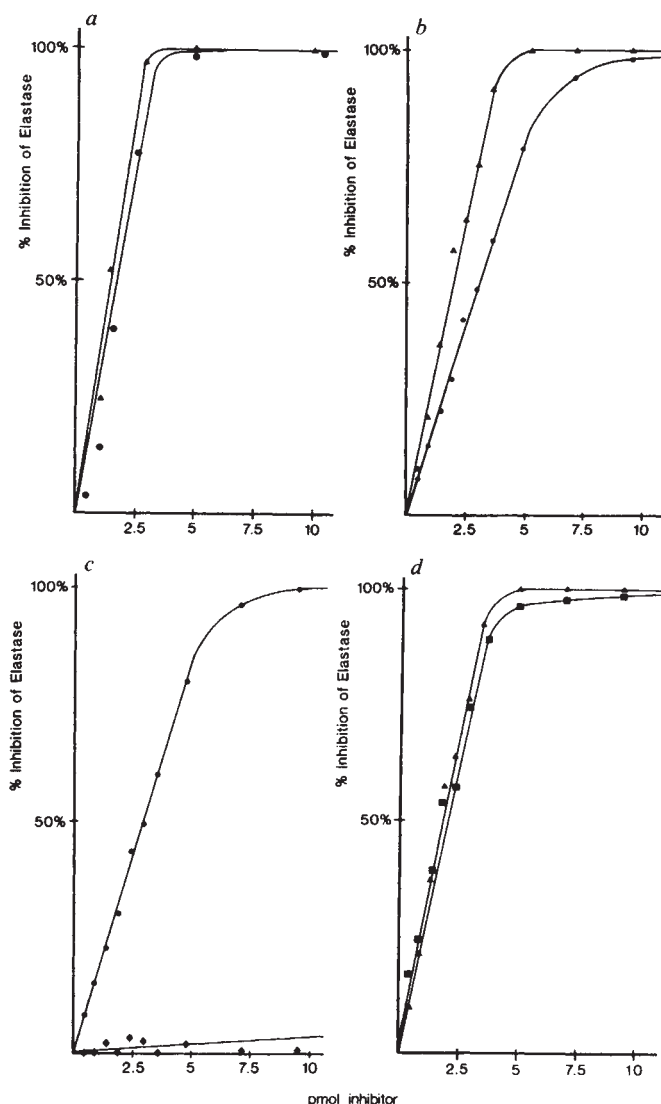


Fig. 3 Inhibition of elastase by α_1 -AT and α_1 -AT(Met 358 $\rightarrow$ Val) and the effect of oxidation by *N*-chlorosuccinimide. a, b, Inhibition of human neutrophil elastase (a) and porcine pancreatic elastase (b) with culture extracts containing α_1 -AT (●) and α_1 -AT(Met 358 $\rightarrow$ Val) (▲). c, Porcine pancreatic elastase with culture extracts containing α_1 -AT untreated (●) and treated with 10 mM *N*-chlorosuccinimide (◆); d, porcine pancreatic elastase with culture extracts containing α_1 -AT (Met 358 $\rightarrow$ Val) untreated (▲) and treated with 10 mM *N*-chlorosuccinimide (■).

Methods: Extracts of *E. coli* cultures, induced as described in Fig. 2 legend, were prepared by sonication and total protein concentration was estimated using a standard assay (BioRad). Concentrations (pmol) of α_1 -AT were calculated based on radial immune diffusion (RID, Calbiochem) assay of α_1 -AT content in the extracts. Elastase assays were performed using either 50 ng human leukocyte elastase (Elastin Products) or 150 ng porcine pancreatic elastase (Sigma) preincubated for 30 min at 23 °C with increasing amounts of cell extracts before addition of 2 mM substrate (*N*-succinyl-Ala-Ala-Pro-Val-*p*-nitroanilide; Calbiochem). Initial reaction rates were measured by monitoring change in absorbance at 410 nm. Oxidation was carried out by preincubation of culture extracts in the presence of 10 mM *N*-chlorosuccinimide (Ega-Chemie).

is both an efficient elastase inhibitor and resistant to oxidation.

A large reduction in elastase inhibitory capacity in the lung may result in extensive damage to the relatively fragile alveolar walls. Although the evidence is controversial^{21,22}, this may occur in individuals who are heavy cigarette smokers, because of direct inactivation of α_1 -AT by oxidants in cigarette smoke and/or oxygen radicals released by stimulated macrophages and neutrophils in the alveolar structures^{13,14}. In this context, α_1 -

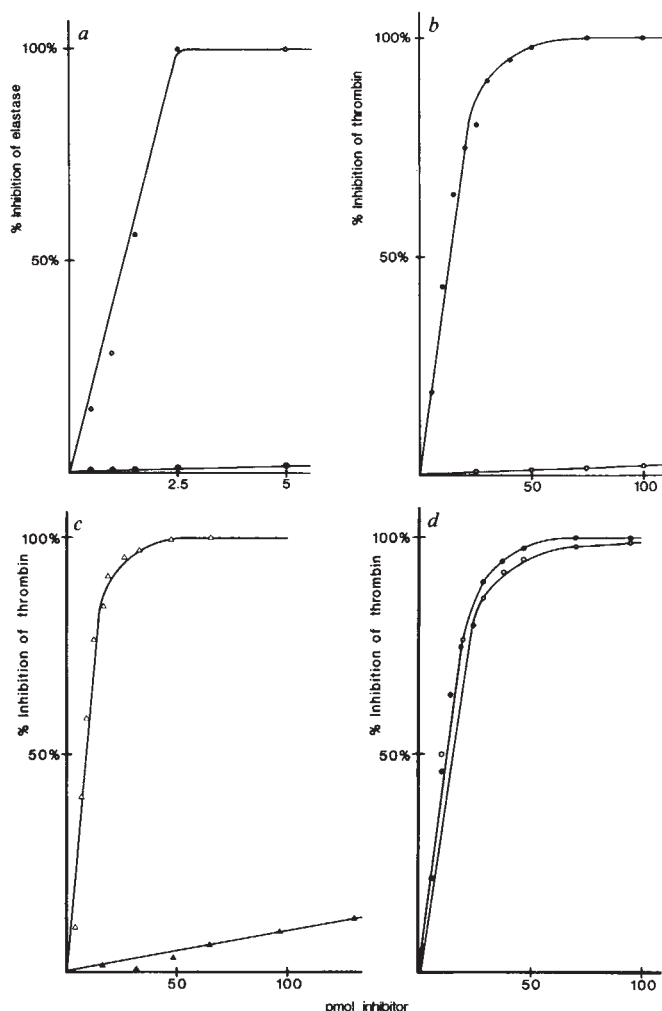


Fig. 4 Inhibition of human neutrophil elastase and thrombin by α_1 -AT and α_1 -AT(Met³⁵⁸→Arg). *a*, Elastase inhibition by *E. coli* extracts containing α_1 -AT (○) and α_1 -AT(Met³⁵⁸→Arg) (●); *b*, thrombin inhibition by extracts containing α_1 -AT (○) and α_1 -AT(Met³⁵⁸→Arg) (●); *c*, thrombin inhibition by purified antithrombin III without (▲) and in the presence of heparin (△); *d*, thrombin inhibition by extracts containing α_1 -AT(Met³⁵⁸→Arg) without (●) and in the presence of heparin (○).

Methods: Human thrombin (0.24 U, Sigma) was preincubated for 20 min at 23 °C with increasing amounts of culture extracts containing inhibitor. After addition of 0.1 mM substrate (Chromozym TH, Boehringer Mannheim), initial reaction rates were obtained by measuring the change in absorbance at 410 nm. Where indicated, preincubation contained 0.15 U heparin (Sigma).

AT(Met³⁵⁸→Val) could be a particularly effective therapeutic agent in the treatment of α_1 -AT deficiency or in protecting the lung in situations of increased oxygen burden. Furthermore, because α_1 -AT(Met³⁵⁸→Val) is resistant to oxidative inactivation, its preparation and storage may be easier than with natural α_1 -AT. This is particularly important because of the large quantities of intravenously administered natural α_1 -AT (~4 g per week) likely to be necessary for effective therapy of α_1 -AT-deficient individuals⁴.

Antielastase activity was not detectable in extracts containing α_1 -AT(Met³⁵⁸→Arg) (Fig. 4*a*). However, in contrast to the unmodified α_1 -AT, this protein showed significant thrombin inhibitory capacity. This observation is of interest because α_1 -AT and antithrombin III share 29% amino acid sequence homology, suggesting that they evolved by divergence from a common ancestral protease inhibitor^{23,24}. Examination of the sequence of the reactive site of antithrombin III (see Fig. 1) shows a P1 Arg residue, the preferred cleavage position of

thrombin. The fact that a single Met→Arg replacement at P1 converts α_1 -AT to an inhibitor of thrombin suggests that the P1 position is crucial in determining the specificity of the inhibitor.

As heparin is known to accelerate plasma antithrombin III inhibition of thrombin, probably by inducing a conformational change in the antithrombin III (refs 25, 26) and/or thrombin itself^{27,28}, we compared the effect of heparin on antithrombin III and α_1 -AT(Met³⁵⁸→Arg). Whereas heparin enhanced the ability of plasma antithrombin III to inhibit thrombin (Fig. 4*c*), this effect was not observed with α_1 -AT(Met³⁵⁸→Arg) (Fig. 4*d*). The level of thrombin inhibition by α_1 -AT(Met³⁵⁸→Arg) was similar to that of heparin-activated antithrombin III. Similar results have been obtained with the natural α_1 -AT (Pittsburgh) variant¹⁸.

The anticoagulant activity of heparin is used clinically in the prevention or treatment of venous thrombosis and pulmonary embolism²⁹. Low-dose administration is used prophylactically in patients at risk, for example following surgery, while high doses are used to treat acute thrombosis. There is, however, an attendant risk of bleeding with the use of heparin, especially at high doses, where the incidence of severe haemorrhage is around 10% (ref. 30). Thus, there is a need for alternative antithrombotic agents that display reduced side effects. In this context, we intend to assess *E. coli*-produced α_1 -AT (Met³⁵⁸→Arg).

The present results illustrate the power of recombinant DNA techniques to manipulate with precision the fine structure of complex proteins. This permits detailed analysis of their mode of action and the design of new molecules for a defined purpose. We show here that site-directed mutagenesis of the reactive site of α_1 -AT can produce proteins with altered or improved properties. These new inhibitors have potential in the therapy of emphysema and thrombotic disease.

We thank Alain Balland for the synthetic oligonucleotides, Dominique Villeval and Francine Jaeger for the DNA sequencing analyses, P. Chambon, P. Kourilsky, J.-P. Cazenave and M. Jaye for many useful suggestions, François Daul for artwork, Paul Tolstoshev and Richard Lathe for reading the manuscript, and Irène Batra and Edith Chambon for secretarial assistance.

Received 17 September; accepted 23 October 1984.

- Carrell, R. W. *et al.* *Nature* **298**, 329–334 (1982).
- Beatty, K. G. thesis, Univ. Georgia (1982).
- Gadek, J. E. & Crystal, R. G. in *The Metabolic Basis of Inherited Diseases* 5th edn (eds Stanbury, J. B., Wyngaarden, J. B., Fredrickson, D. S., Goldstein, J. L. & Brown, M. S.) 1450–1467 (McGraw-Hill, New York, 1981).
- Gadek, J. E., Klein, H. G., Holland, P. V. & Crystal, R. G. *J. clin. Invest.* **68**, 1158–1165 (1981).
- Eriksson, S. *Acta med. scand.* **175**, 197–205 (1964).
- Kuhn, C. & Senior, R. M. *Lung* **155**, 185–197 (1978).
- Gadek, J. E. *et al.* *Bull. Eur. Physiopath. Resp.* **16**, Suppl., 27–40 (1980).
- Kueppers, F. in *Genetic Determinants of Pulmonary Disease* (ed. Litwin, S. D.) 23–74 (Dekker, New York, 1978).
- Morse, J. O. *New Engl. J. Med.* **299**, 1045–1099 (1978).
- Larsson, C. *Acta med. scand.* **204**, 345–351 (1978).
- Carp, H. & Janoff, A. *Am. Rev. resp. Dis.* **118**, 617–621 (1978).
- Abrams, W. R., Eliraz, A., Kimbel, P. & Weinbaum, G. *Exp. Lung Res.* **1**, 211–223 (1981).
- Gadek, J. E., Fells, G. A. & Crystal, R. G. *Science* **206**, 1315–1316 (1979).
- Carp, H., Miller, F., Hoidal, J. R. & Janoff, A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2041–2045 (1982).
- Johnson, D. & Travis, J. *J. biol. Chem.* **254**, 4022–4026 (1979).
- Beatty, K., Bieth, J. & Travis, J. *J. biol. Chem.* **255**, 3931–3934 (1980).
- Lewis, J. H., Iammarino, R. M., Spero, J. A. & Hasiba, U. *Blood* **51**, 129–137 (1978).
- Owen, M. C., Brennan, S. O., Lewis, J. H. & Carrell, R. W. *New Engl. J. Med.* **309**, 694–698 (1983).
- McRae, B., Nakajima, K., Travis, J. & Powers, J. C. *Biochemistry* **19**, 3973–3978 (1980).
- Courtney, M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **81**, 669–673 (1984).
- Stone, P. J. *et al.* *Science* **221**, 1187–1189 (1983).
- Boudier, C., Pelletier, A., Pauli, G. & Bieth, J. G. *Clin. chim. Acta* **132**, 309–315 (1983).
- Carrell, R. W., Boswell, D. R., Brennan, S. O. & Owen, M. C. *Biochem. biophys. Res. Commun.* **93**, 399–402 (1980).
- Hunt, L. T. & Dayhoff, M. O. *Biochem. biophys. Res. Commun.* **95**, 864–871 (1980).
- Rosenberg, R. D. & Damus, P. S. *J. biol. Chem.* **248**, 6490–6505 (1973).
- Villanueva, G. B. & Danishefsky, I. *Biochemistry* **18**, 810–817 (1979).
- Griffith, M. J. *J. biol. Chem.* **257**, 7360–7365 (1982).
- Nesheim, M. E. *J. biol. Chem.* **258**, 14708–14717 (1983).
- Cazenave, J.-P., Wiesel, M.-L. & Hemmendinger, S. *Ag. Actions* **15**, 24–49 (1984).
- Mant, M. J. *et al.* *Lancet* **i**, 1133–1135 (1977).
- Tessier, L.-H. *et al.* *Nucleic Acids Res.* **12**, 7663–7675 (1984).
- Winter, G., Fersht, A. R., Wilkinson, A. J., Zoller, M. & Smith, M. *Nature* **299**, 756–758 (1982).
- Zoller, M. & Smith, M. *Nucleic Acids Res.* **10**, 6487–6500 (1982).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Lathe, R., Lecocq, J.-P. & Everett, R. in *Genetic Engineering* Vol. 4 (ed. Williamson, R.) 1–56 (Academic, London, 1983).

Site-directed mutagenesis of cytochrome *c* shows that an invariant Phe is not essential for function

Gary J. Pielak, A. Grant Mauk and Michael Smith

Department of Biochemistry, Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5

Phenylalanine 87 of yeast iso-1-cytochrome *c* (Phe 82 in horse heart and bonito) is phylogenetically conserved¹ and occurs near the surface of the protein². It has been suggested that this residue is directly involved in electron transfer between cytochrome *c* and cytochrome *c* peroxidase (CCP)³ and may also control the polarity of the haem environment⁴. Because Phe residues are not susceptible to chemical modification, no direct means of studying the functional role of Phe 87 has been available, so we have chosen Phe 87 as our initial target here to test the feasibility of using site-directed mutagenesis⁵ as a means of studying structure-function relationships in cytochrome *c*. We have changed the codon for Phe 87 to that of either a Ser, a Tyr or a Gly. The mutated genes have been introduced⁶ into a yeast strain⁷ lacking both isozymes of cytochrome *c*. Unlike the recipient strain, transformants grow on a non-fermentable carbon source, indicating that the mutant proteins can reduce cytochrome oxidase⁷. The purified mutant proteins are similar to wild type with respect to their visible spectra⁸, 20–70% as active as wild-type protein in the CCP assay⁹, and their reduction potentials are lowered by as much as 50 mV. Thus Phe 87 is not essential for cytochrome *c* to transfer electrons but is involved in determining the reduction potential.

For *in vitro* mutagenesis, a 2.5-kilobase (kb) *Bam*HI-*Hind*III fragment of established DNA sequence (refs 10–12 and D. W. Leung and M.S., unpublished results) containing the iso-1-cytochrome *c* gene (*CYC1*) locus has been inserted into M13mp8 (ref. 13). The recombinant vector was mutated by the procedure of Zoller and Smith⁵ or a simpler, two-primer method as shown in Fig. 1 (ref. 14 and M. J. Zoller and M.S., unpublished) using the synthetic oligodeoxyribonucleotides shown in Fig. 2. Mutants of *CYC1* in M13mp8 were detected by probing with ³²P-labelled mutagenic oligodeoxyribonucleotides using empirically established stringent hybridization^{5,15,16} of either phage dot-blots or plaque-lifts¹⁷. The efficiency of mutagenesis, as determined by hybridization screening, was 8%, 13% and 15%, respectively, for mutants containing Ser 87, Tyr 87 and Gly 87. The mutations were confirmed by DNA sequence determination using the chain terminator method (Fig. 3)^{18,19}.

The double-stranded 2.5-kb fragment for the wild type and for each of the mutants was isolated and cloned into the yeast shuttle vector YEp213 (ref. 6). The recombinant plasmids were isolated via *Escherichia coli* MM294 (ref. 20) and used to transform the *Saccharomyces cerevisiae* recipient strain GM-3C-2 by the LiCl method²¹. Strain GM-3C-2 carried *leu2* and *trp1* markers and a deletion mutation (*cyc1-1*) and a point mutation (*cyp3-1*) which eliminate and inactivate the genes for iso-1- and iso-2-cytochrome *c*, respectively⁷. Therefore GM-3C-2 cannot grow on glycerol or lactate. To avoid selection for genomic revertants in the genes for cytochrome *c*, transformants were first selected for leucine prototrophy. *Leu*⁺ transformants containing the wild-type iso-1-cytochrome *c* and the three mutants all grow on glycerol or lactate as sole carbon source, indicating that the gene product can reduce cytochrome oxidase in all four cases. The plasmids were recovered from the transformants via MM294²⁰, and the double-stranded DNA was sequenced to re-confirm the presence of the mutations. Curing the four yeast transformants of their respective plasmids causes their phenotypes to revert to that of GM-3C-2, which indicates that no reversion in *cyc1-1* or *cyp3-1* had occurred.

Cells containing the mutated genes were grown to saturation (3×10^8 cells per ml) on glycerol, and the cytochromes *c* were

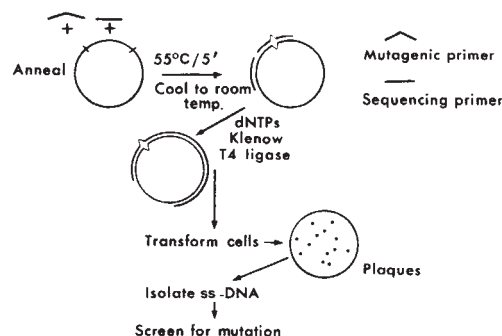


Fig. 1 Schematic representation of 2-primer mutagenesis. The primers are not drawn to scale. The 5'-phosphorylated (15 pmol) mutagenic oligodeoxyribonucleotide primer and the M13 universal primer (7.5 pmol) were annealed to the template (0.5 pmol) in a total of 12 μl and then elongated (17.5 h) and ligated⁵. The reaction mixture was diluted 1 to 10, and 1 μl was used to transform competent *E. coli* JM101 cells. This resulted in about 200 clear plaques. A plaque lift¹⁷ and subsequent screening with the 5'-³²P-labelled⁵ mutagenic oligodeoxyribonucleotide revealed an efficiency of 15%. Several positive plaques were picked, plated out and DNA from individual plaques was screened using dot-blot hybridization⁵.

```

3'          5'
GGGTGGTTTCCGGTA  Template

5'          3'
CCCACCAAAGGCCAT  Phe (wild type)

CCCACCAGAGGCCAT  Ser

CCCACCATAGGCCAT  Tyr

CCCACCACCGGCCAT  Gly

```

Fig. 2 The oligodeoxyribonucleotides used to effect site-directed mutagenesis at codon 87 of the sense strand of yeast iso-1-cytochrome *c* gene. They were synthesized on glass beads using the chemistry described in ref. 26. The preparation of the intermediate diisopropyl phosphoramidites and the protocols for manual synthesis, deprotection and purification are described in ref. 27.

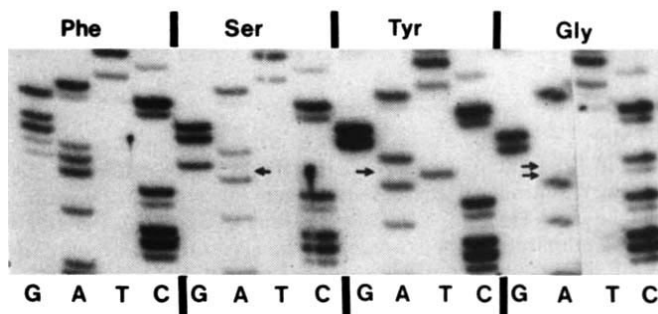


Fig. 3 Comparison of wild-type (phe) and mutant sequences (ser, tyr and gly) by chain-terminator sequencing^{18,19}. The site of each mutation is indicated by an arrow. The sequence is that of the anti-sense strand of *CYC1* (ref. 11). Thus, the codon number increases from top to bottom and the triplet of A residues in the wild-type sequence represents the complement of the Phe 87 codon as shown in Fig. 1. The sequencing primer was 5'-CATGATATCGASCAAAG-3' which primes 90 base pairs upstream from the A triplet. This 8% polyacrylamide gel was electrophoresed for 2 h at 1,500 V.

Table 1 Amino acid analysis of purified wild-type and mutant cytochromes *c*

Amino acid	Sample (average residues per molecule)				
	Wild type ¹²	Phe 87	Ser 87	Tyr 87	Gly 87
Ala	7	7.3 (0.1, 16)*	7.5 (0.3, 15)	7.3 (0.2, 12)	7.3 (0.1, 14)
Arg	3	3.5 (0.6, 16)	3.5 (0.4, 16)	3.6 (0.1, 12)	3.5 (0.2, 14)
Asx	11	11.2 (0.4, 16)	11.3 (0.3, 15)	11.2 (0.4, 12)	11.2 (0.3, 14)
Cys	3	2.4 (0.5, 6)	2.7 (0.6, 6)	2.9 (0.5, 3)	2.9 (0.4, 4)
Glx	9	9.4 (0.4, 16)	9.2 (0.2, 15)	9.3 (0.3, 12)	9.4 (0.3, 14)
Gly	12	12.0 (0.8, 16)	12.3 (0.4, 15)	11.8 (0.4, 12)	12.9 (0.2, 14)
His	4	4.0 (0.2, 16)	4.0 (0.2, 15)	4.1 (0.1, 12)	4.1 (0.2, 14)
Ile	4	3.5 (0.3, 16)	3.4 (0.4, 15)	3.8 (0.3, 12)	3.5 (0.3, 14)
Leu	8	7.9 (0.4, 16)	7.7 (0.2, 15)	7.9 (0.2, 12)	8.0 (0.2, 14)
Lys	15	14.9 (1.2, 16)	15.2 (1.0, 15)	14.8 (0.6, 12)	15.1 (1.2, 14)
Met	2	1.7 (0.4, 8)	2.0 (0.2, 8)	2.2 (0.3, 8)	1.9 (0.2, 12)
Phe	4	3.8 (0.5, 16)	2.8 (0.3, 15)	2.9 (0.1, 12)	3.0 (0.3, 14)
Pro	4	4.3 (0.4, 16)	4.7 (0.6, 15)	5.0 (0.8, 12)	4.3 (0.4, 14)
Ser	4	3.7 (0.3, 12)	4.7 (0.3, 15)	3.8 (0.3, 12)	3.9 (0.4, 14)
Thr	8	7.6 (0.4, 16)	7.6 (0.2, 15)	7.6 (0.2, 12)	7.8 (0.2, 14)
Trp	1	1	1	1	1
Tyr	5	4.9 (0.3, 10)	4.6 (0.3, 9)	6.0 (0.1, 8)	4.4 (0.2, 10)
Val	3	3.0 (0.2, 16)	2.7 (0.3, 15)	2.8 (0.5, 12)	3.0 (0.1, 14)
tLys†	1	0.9 (0.2, 12)	1.1 (0.2, 15)	1.1 (0.1, 12)	1.1 (0.2, 14)

Hydrolyses were carried out in 6 M HCl at 112 °C and 0.1 mm Hg after purging with N₂. Cys was determined as cysteic acid after performic acid oxidation. Met was quantified as itself or as methionine sulphone. Trp content was estimated from inspection of the near-UV spectrum. Calculations were based on the total no. of nmoles of amino acids detected in a chromatogram over the total no. of residues in the protein.¹⁰

* In parenthesis: standard deviation, no. of determinations.

† tLys, N⁶-trimethyl lysine²⁵.

purified by a procedure²² that readily separates the two isozymes and their derivatives. No band for iso-2-cytochrome *c* was observed in preparations from any of the transformants, suggesting that a reversion had not taken place. More than 3 mg of cytochrome *c* were obtained per litre of culture. The results of amino acid analyses (Table 1) indicate that the amino acid content of the mutant proteins is identical to that of the wild-type protein except that each mutant lacks one Phe residue and has gained, depending on the mutant, either a Ser, Tyr or Gly residue.

The peak positions in the visible (750–325 nm) spectra of both the oxidized and reduced mutant proteins are within 2.5 nm and the relative intensities of the maxima are within 10% of the values for the wild-type protein⁸. That these mutations have little effect on the visible spectra is not surprising, as Phe 87 is more than 5 Å away from the haem chromophore. Ligation of Met to the iron atom in all of the mutants is apparent from the presence of the band at 695 nm in each case. The near-UV difference spectra (oxidized mutant proteins minus wild type) all show a feature near 290 nm, probably indicating movement of the sole Trp residue. The difference spectrum for the Tyr 87 mutant between 265 and 280 nm is very nearly the spectrum of free tyrosine, as would be expected when Phe is replaced by Tyr.

The oxidation of reduced cytochrome *c* with 0.6 nM CCP (purified from commercial yeast as described in ref. 23) and 170 μM H₂O₂ in 0.2 M Tris-HCl, pH 7.0 was followed at 550 nm⁹. At a cytochrome *c* concentration of 5 μM, the Ser, Tyr and Gly-87 variants are approximately 70%, 30% and 20% as active as the wild-type protein, respectively. The activities are reproducible to within 2% and cannot be accounted for by autocatalytic oxidation of the mutant cytochromes *c* because the ratio of autocatalytic to enzymatic initial velocities is always greater than 12 for both the wild-type and mutant proteins. Preliminary electrochemical experiments using an optically transparent thin-layer electrode²⁴ indicate that the reduction potentials of the Ser 87 and Gly 87 proteins are equal and 50 mV lower than the Tyr 87 and the wild-type protein. This is as expected if, as implied by Kassner⁴, the function of Phe 87 is to control the polarity of the haem environment by excluding water. It follows that replacement of Phe 87 with Gly or Ser allows greater solvent access to the haem. The Nernst slopes observed with the wild-type and mutant proteins provide further evidence of their homogeneity.

These results demonstrate that Phe 87 is not an absolute requirement for electron transfer between cytochrome *c* and CCP or cytochrome oxidase, although all three mutations reduce the rate of reaction with CCP. The relative activities reported here do not correlate with the effects of the mutations on the potentials alone, so alteration in protein-protein interactions is likely to be an important consideration. As with any kinetics experiment, one cannot rule out the possibility that Phe 87 functions as suggested³ in the wild-type protein, and that its replacement simply reveals alternative, less efficient pathways for electron transfer. Further studies, including a detailed steady-state kinetics analysis, are planned to provide insight into the mechanistic basis of these observations.

We thank Dr Mark J. Zoller, Dr Marcia Mauk, and Ms Gerry Weinmaster for helpful discussions. Dr Helen Whitely and Mr Mas Takahara for use of the fermentor in the Department of Microbiology at the University of Washington and Mr Don McKenzie for operating the amino acid analyser. This research was supported by grants from the MRC of Canada. G.J.P. is a recipient of a postdoctoral fellowship from the US NIH and M.S. is a career investigator of the MRC of Canada.

Received 20 March; accepted 21 September 1984.

- Dickerson, R. E. & Timkovich, R. *Enzymes* **11**, 397–547 (1975).
- Takano, T. & Dickerson, R. E. *J. molec. Biol.* **153**, 79–94 (1981).
- Poulos, T. L. & Kraut, J. *J. biol. Chem.* **255**, 10322–10330 (1980).
- Kassner, R. J. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2263–2267 (1972).
- Zoller, M. J. & Smith, M. *Meth. Enzym.* **100**, 468–500 (1983).
- Hicks, J. B., Strathern, J. N., Klar, A. J. S. & Dellaporta, S. L. in *Genetic Engineering: Principles and Methods* Vol. 4 (eds Setlow, J. K. & Hollaender, A.) 219–248 (Plenum, New York, 1982).
- Faye, G., Leung, D. W., Tatchell, K., Hall, B. D. & Smith, M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2258–2262 (1981).
- Margoliash, E. & Frohwirt, N. *Biochem. J.* **71**, 570–572 (1959).
- Kang, C. H., Ferguson-Miller, S. & Margoliash, E. *J. biol. Chem.* **252**, 919–926 (1977).
- Montgomery, D. L., Hall, B. D., Gillam, S. & Smith, M. *Cell* **14**, 673–680 (1978).
- Smith, M. *et al. Cell* **16**, 753–761 (1979).
- Lederer, F., Simon, A. M. & Verdier, J. *Biochem. biophys. Res. Commun.* **47**, 55–58 (1972).
- Messing, J. *Meth. Enzym.* **101**, 20–78 (1983).
- Norris, K., Norris, F., Christiansen, L. & Fill, N. *Nucleic Acids Res.* **11**, 5103–5112 (1983).
- Astell, C. R., Doel, M. T., Jahnke, P. A. & Smith, M. *Biochemistry* **12**, 5068–5074 (1973).
- Gillam, S., Waterman, K. & Smith, M. *Nucleic Acids Res.* **2**, 625–634 (1975).
- Woo, S. L. C. *Meth. Enzym.* **68**, 389–395 (1979).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161–178 (1980).
- Hanahan, D. *J. molec. Biol.* **166**, 557–580 (1983).
- Ito, H., Fukuda, Y., Murata, K. & Kimura, A. *J. Bact.* **153**, 163–168 (1983).
- Sherman, F. *et al. J. biol. Chem.* **243**, 5446–5456 (1968).

23. Nelson, C. E., Sitzman, E. V., Kang, C. H. & Margoliash, E. *Analyt. Biochem.* **83**, 622-631 (1977).
24. Reid, L. S., Taniguchi, V. T., Gray, H. B. & Mauk, A. G. *J. Am. chem. Soc.* **104**, 7516-7519 (1982).
25. Delange, R. J., Glazer, A. N. & Smith, E. L. *J. biol. Chem.* **245**, 3325-3327 (1970).
26. Adams, S. P., Kavka, K. S., Wykes, E. J., Holder, S. B. & Galluppi, G. R. *J. Am. chem. Soc.* **105**, 661-663 (1983).
27. Atkinson, T. & Smith, M. in *Oligonucleotide Synthesis: A Practical Approach* (ed. Gait, M.) (IRL, Washington DC, in the press).

DNA bending induced by cruciform formation

Gerald W. Gough & David M. J. Lilley

Department of Biochemistry, University of Dundee,
Dundee DD1 4HN, UK

Cruciform structures^{1,2} in DNA are of considerable interest, both as extreme examples of sequence-dependent structural heterogeneity and as models for four-way junctions such as the Holliday junction³ of homologous genetic recombination. Cruciforms are of lower thermodynamic stability than regular duplex DNA, and have been observed only in negatively supercoiled molecules⁴⁻⁶, where the unfavourable free energy of formation is offset by the topological relaxation of the torsionally stressed molecule. From an experimental viewpoint this can be a disadvantage, as cruciform structures can be studied only in relatively large supercoiled DNA circles, and are destabilized when a break is introduced at any point. We therefore set out to construct a pseudo-cruciform junction—by generating heteroduplex formation between two inverted repeat sequences. Stereochemically, this should closely resemble a true cruciform but remain stable in a linear DNA fragment. We have now created such a junction and find that it has the expected sensitivities to endonucleases. These DNA fragments exhibit extremely anomalous gel electrophoretic mobility, the extent of which depends on the relative position of the pseudo-cruciform along the length of the molecule. Our results are very similar to those obtained by Wu and Crothers⁷ using kinetoplast DNA, and we conclude that the pseudo-cruciform junction introduces a bend in the linear DNA molecule.

Cruciform-like structures that are stable in linear DNA molecules are desirable for several reasons. First, the four-way junction is very interesting both intrinsically and as a model for the Holliday junction. The stereochemistry of this junction can be probed in supercoil-stabilized cruciforms, as illustrated by studies with the resolving enzyme T4 endonuclease VII (refs 8, 9). However, more detailed analysis requires a stable junction. Second, the usefulness of immunochemical techniques for studying alternative DNA structures¹⁰ suggests that a similar approach for cruciforms might be worthwhile. Attempts to raise antibodies to a supercoiled DNA species are likely to be frustrated by the action of nucleases in serum, however, and an immunogen that is independent of topology is therefore required. Third, we wish to investigate the possible existence of cellular proteins that specifically bind four-way junctions, and a DNA fragment containing a pseudo-cruciform junction should provide a suitable tool for the search.

Four-way junctions have been obtained *in vitro* in several ways, including cell-free recombination¹¹, isolation of strand-exchange forms of 2 μ DNA from yeast¹² and hybridization of phage λ *att-B* and *att-P* sites⁸, while chemical construction of a stable four-way junction has been reported by Kallenbach *et al.*¹³. The latter species may be very useful for crystallographic and spectroscopic studies, but their small size hampers enzymatic investigation. The basic principle of the approach adopted by us (Fig. 1) involves heteroduplex formation between two inverted repeats of unrelated sequences which are flanked by identical sequences. In these circumstances the sequences hybridize perfectly except for the region of 2-fold symmetry, where duplex formation is prevented by the absence of homology. Rather than remaining as a simple unpaired bubble,

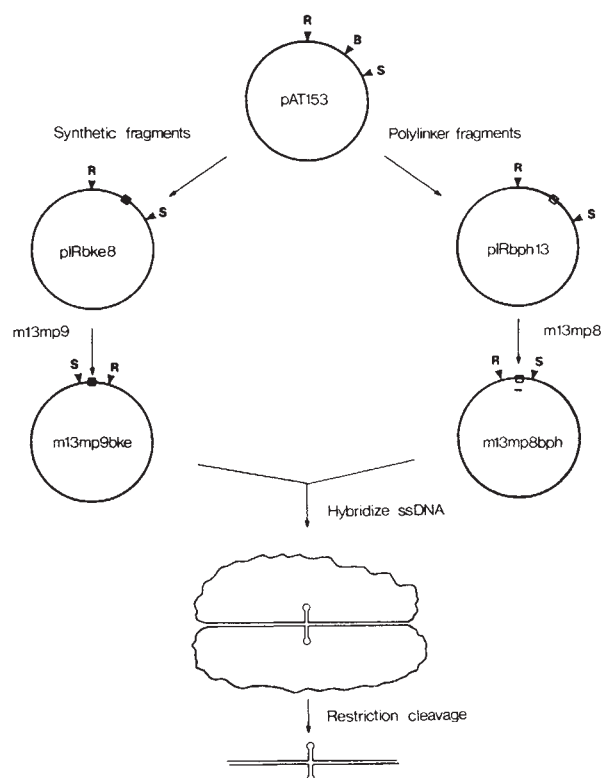


Fig. 1 Construction of a pseudo-cruciform junction by hybridization of two inverted repeats. Plasmid molecules containing inverted repeats of unrelated sequence were constructed by inserting into plasmid pAT153 (ref. 21) oligonucleotide fragments (the construction of pIRbke8 is described in ref. 14) or an inverted dimer of the *Bam*HI-*Hind*III fragment of the polylinker from phage m13mp8 (ref. 13). These inverted repeats, each residing in the *Bam*HI site of the parent vector, are shown as filled and open boxes, respectively. R, *Eco*RI site; B, *Bam*HI site; S, *Sal*I site. The ~680-bp *Eco*RI-*Sal*I fragments were excised from pIRbke8 and pIRbph13 (restriction sites contained within the inverted repeats necessitated partial digests), subcloned into the phages m13mp9 and m13mp8¹⁵, respectively, and propagated in *Escherichia coli* JM103. M13 phage was recovered from cellular supernatants by precipitation with 20% polyethylene glycol, 2.5 M NaCl on ice. Redissolved phage was extracted with phenol and the single-stranded DNA precipitated with ethanol. In a typical hybridization, 30 μ g of each DNA species was incubated at 65 °C for 3 h in 0.15 M NaCl, 0.015 M Na-citrate in a total volume of 50 μ l. The hybridized DNA was then ethanol precipitated, redissolved in water and cleaved by restriction enzymes as directed by the manufacturer (New England Biolabs or Bethesda Research Labs). DNA fragments containing a pseudo-cruciform were purified by gel electrophoresis on 5% polyacrylamide, excision of the appropriate ethidium bromide stained bands and electroelution onto DEAE cellulose (Whatman). The fragments were recovered from the resin by elution with 2M NaCl, and ethanol-precipitated.

however, the two inverted repeats can undergo intra-strand base pairing. Although stereochemically identical to a cruciform, the structure obtained is not a true cruciform because the nucleotide sequences of the two arms are unrelated by symmetry. Unlike a true cruciform, this structure cannot be converted to a perfect duplex, and will therefore remain stable in an unstressed molecule.

We used plasmid pAT153 as the parent cloning vector and generated two recombinant molecules by constructing inverted repeats at the unique *Bam*HI site. Plasmid pIRbke8 contains a 32-base pair (bp) *Bam*HI-*Eco*RI-*Bam*HI inverted repeat constructed by cloning two synthetic oligonucleotides¹⁴, while pIRbph13 contains a 46-bp *Bam*HI-*Hind*III-*Bam*HI inverted repeat created by self-ligation of the *Bam*HI-*Hind*III region of the polylinker of phage m13mp8. Both species undergo

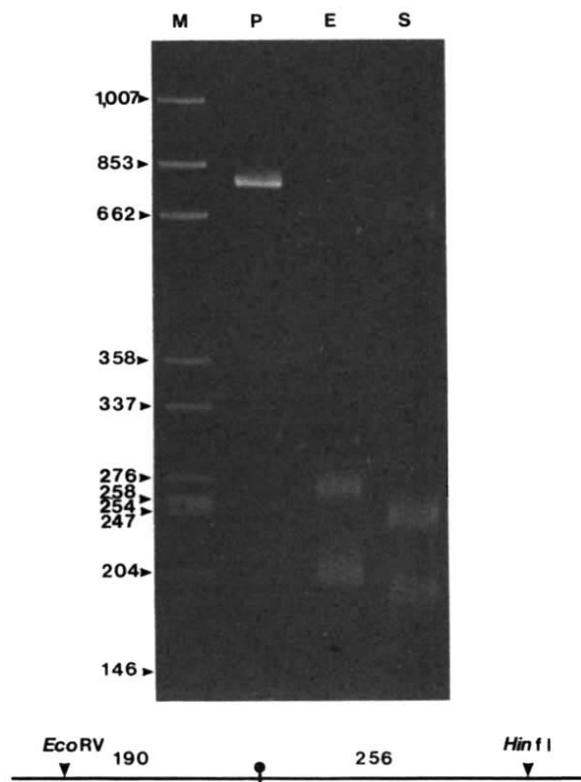


Fig. 2 A pseudo-cruciform is sensitive to S_1 nuclease and T4 endonuclease VII. Gel-purified *EcoRV-HinI* fragment was incubated with 5 U S_1 nuclease (BRL) for 12 h at 4°C and loaded directly onto the gel. The fragment was separately incubated with 500 U T4 endonuclease VII in 50 mM Tris pH 8.0, 10 mM $MgCl_2$, 10 mM β -mercaptoethanol, 250 $\mu g\ ml^{-1}$ bovine serum albumin for 3 h at 37°C, and again loaded directly onto the gel. 5% polyacrylamide gel electrophoresis was performed in 90 mM Tris pH 8.3, 90 mM borate, 10 mM EDTA at ambient temperature for 16 h. Tracks (left to right) contained: M, marker DNA fragments (size in base pairs) derived by cleavage of phage Φ X174 RF DNA with *AluI*; P, *EcoRV-HinI* pseudo-cruciform fragment; E, fragment after cleavage by T4 endonuclease VII; S, fragment after cleavage by S_1 nuclease. The diagram shows the positions of the pseudo-cruciform and the *EcoRV* and *HinI* restriction enzyme target. Fragment lengths are calculated to the bases of the pseudo-cruciform arms.

cruciform formation in the supercoiled plasmids. Heteroduplex formation between these molecules has been achieved, but the results are complicated by the mixture of two possible heteroduplex and two homoduplex products. We have simplified this situation to a single heteroduplex by introducing a further sub-cloning step in a single-strand DNA phage. The ~680-bp *EcoRI-SalI* fragments of pIRbke8 and pIRbph13 were purified and cloned into phages m13mp9 and m13mp8 (ref. 15), respectively. Because the cloning sites of these two vectors are in opposite orientations, the single-stranded DNA of the phage contains *EcoRI-SalI* inserts that are complementary and can therefore be hybridized. As the phage always encapsidates the same (+)-strand, only one heteroduplex product can result. Just as for the above situation, the hybrid may form along the entire length of the *EcoRI-SalI* region (the rest of the phage DNA remains unhybridized as it is all the same strand), except for the region of the inverted repeat, where a pseudo-cruciform results. The many restriction enzyme target sites present along this DNA enable the fragment containing the pseudo-cruciform to be cleaved out and purified.

DNA fragments containing this pseudo-cruciform can be cleaved by nucleases to which cruciforms are sensitive. Figure 2 shows the cleavage of the *EcoRV-HinI* fragment by S_1 nuclease and T4 endonuclease VII. In each case the pseudo-

cruciform fragment is completely digested and two 'half-fragment' bands are generated. With S_1 nuclease, these are 244 and 187 $\pm$ 5 bp in length, consistent with the removal of the entire pseudo-cruciform. This is to be expected, as cleavage of the cruciform loops and separation of the half-molecules leaves the remaining 'stem' DNA unable to form base pairs, and therefore sensitive to S_1 nuclease. In contrast, T4 endonuclease VII cleaves diagonally across the four-way junction^{8,9}, producing fragments which should be longer by a single stem-length, and the measured lengths of 263 and 199 $\pm$ 5 bp are in good agreement with this. The resolvase is extremely structure-specific, and probably provides the strongest evidence for the existence of the pseudo-cruciform. Cleavage by S_1 nuclease also generates a small proportion of material migrating at ~600 bp. We have not investigated this species in detail, but suspect that it may result from S_1 nuclease cleavage on one cruciform arm only, possibly creating a flexible joint. This species is not produced by T4 endonuclease VII, which always introduces simultaneous double cleavages. The arms on the polylinker side of the pseudo-cruciform are sensitive to *PstI* and *SalI* (detected from the 5'-termini generated), but the hybrids are totally refractory to cleavage by *EcoRI* and *HindIII*, the sites for which should be present in the loops. Taken together, these data provide strong evidence for the generation of a cruciform-like four-way junction.

Comparison of the migration of the pseudo-cruciform fragment with the marker DNA fragments shows (Fig. 2) that its mobility in 5% polyacrylamide is anomalously slow, unlike any of the component molecules. Table 1 shows the apparent mobility of the ~400-bp *EcoRV-SphI* fragment as a function of the gel matrix. Whereas, in agarose, the migration is only slightly anomalous, in 5% polyacrylamide the fragment has the apparent mobility of a 660-bp fragment, and the anomaly increases with polyacrylamide concentration. This behaviour is totally independent of the strength of the electric field, but is sensitive to the gel temperature during the electrophoresis run. The effect is so striking that we considered possible causes other than the frictional drag of the pseudo-cruciform arms. We had suspected previously that such a junction might affect the geometry of the entire molecule, perhaps in a way reminiscent of the bending sequence of kinetoplast DNA. Following Wu and Crothers⁷, we therefore investigated whether the relative position of the pseudo-cruciform on the fragment affected the migration anomaly. The nature of the procedure for obtaining the DNA prevented us from using rigorous cyclic permutations of fragments. Nevertheless, by choosing various combinations of restriction enzymes (Fig. 3a), it was possible to prepare a series of fragments in which the position of the pseudo-cruciform varied from a central location to sites close to either end. By examining the migration of some typical fragments (Fig. 3b) and plotting apparent size divided by calculated length as a function of relative pseudo-cruciform position (Fig. 3c), we found that the degree of migration anomaly is profoundly affected by pseudo-cruciform position. When the pseudo-cruciform is central, its retarding effect on migration is much greater than when it is at either end, and intermediate positions have correspondingly intermediate relative mobilities, giving a smooth

Table 1 Anomalous electrophoretic migration of a pseudo-cruciform fragment as a function of gel concentration

Gel matrix	Apparent sizes (bp)
2% agarose	445
5% polyacrylamide	660
8% polyacrylamide	850
12% polyacrylamide	1,750

The apparent size of the 400 bp *EcoRV-SphI* pseudo-cruciform fragment was estimated in each gel matrix by interpolation from normal linear marker fragments of DNA.

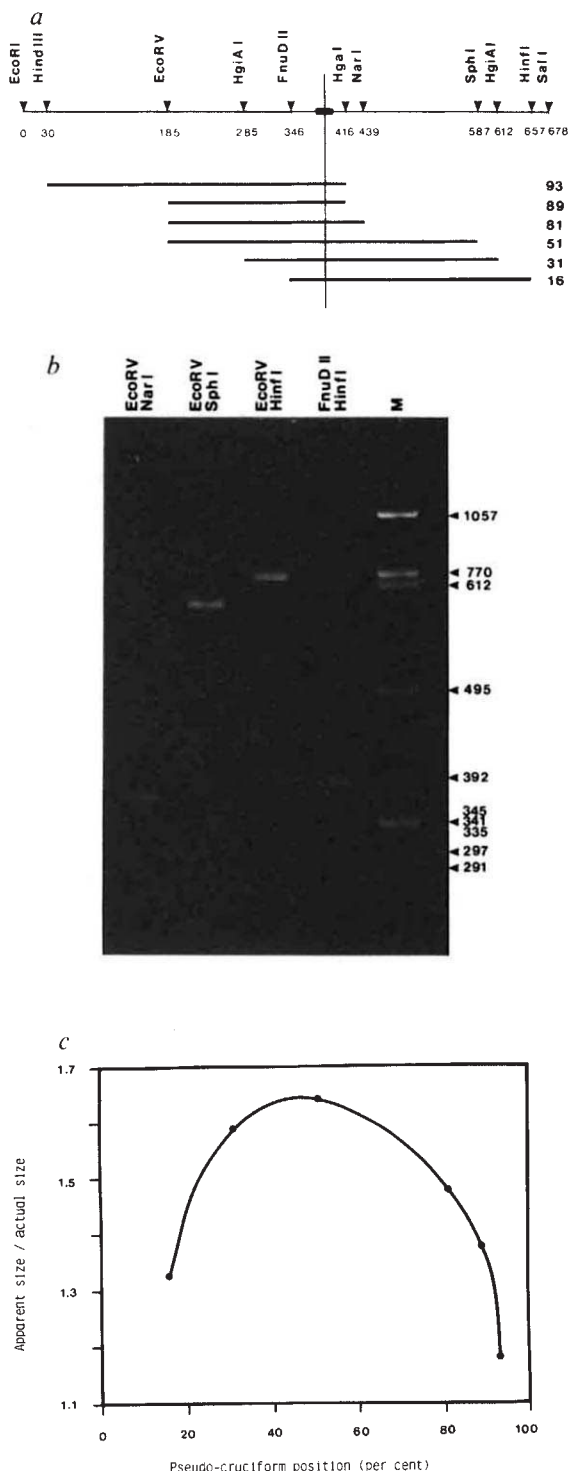


Fig. 3 Anomalous electrophoretic migration is a sensitive function of pseudo-cruciform position. **a**, Restriction map of the EcoRI-SalI fragment containing the pseudo-cruciform (box). Various restriction fragments used in the mobility measurements are indicated below the map. The number on the right of each fragment is the position of the pseudo-cruciform calculated as a percentage of the total length. **b**, 5% polyacrylamide gel electrophoresis of typical pseudo-cruciform fragments. The track labelled M contains a HincII digest of phage Φ X174 RF DNA, to provide conventional linear marker fragments, in order to measure apparent sizes of pseudo-cruciform fragments. We note that the 612-bp fragment of this digest migrates somewhat anomalously. For accurate size calibration, we have used a combination of HincII, AluI, HaeIII and HinfI digests of Φ X174 RF DNA. **c**, Plot of ratios of apparent size to calculated size against the relative position of the pseudo-cruciform junction within the fragments.

dome-shaped curve. This result would not be expected if the retardation due to the cruciform arms was caused by frictional effects. Moreover, if one arm is shortened by restriction enzyme cleavage, a relatively small reduction in migration anomaly is seen.

We conclude that the pseudo-cruciform perturbs the geometry of the entire molecule by introducing a bend or possibly a point of flexibility. This should exert a greater influence on the overall shape of the molecule the closer to the centre it occurs. The most symmetrical arrangement for four DNA helices is tetrahedral, and we have previously successfully used this basic assumption to analyse T4 endonuclease VII cleavage of two short cruciforms⁹. A tetrahedral arrangement would introduce a bend of 109° into the fragments analysed by gel electrophoresis. It is unclear how such a junction could be flexible, but we cannot completely exclude this possibility.

We therefore believe that we have identified a structural perturbation in a DNA molecule which introduces a pronounced bend; there is both practical^{7,16} and theoretical¹⁷⁻²⁰ evidence for this in other systems. Our pseudo-cruciform fragments will be useful as an experimental model for DNA bending, as well as for the study of four-way junctions. Furthermore, immunological experiments are in progress.

We thank Dr B  rries Kemper for provision of T4 endonuclease VII, and the MRC and the Royal Society for financial support.

Received 7 September; accepted 17 October 1984.

1. Platt, J. R. *Proc. natn. Acad. Sci. U.S.A.* **41**, 181-183 (1955).
2. Gierer, A. *Nature* **212**, 1480-1481 (1966).
3. Holliday, R. *Genet. Res.* **5**, 282-304 (1964).
4. Gellert, M., Mizuuchi, K., O'Dea, M. H., Ohmori, H. & Tomizawa, J. *Cold Spring Harb. Symp. quant. Biol.* **43**, 35-40 (1979).
5. Lilley, D. M. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6468-6472 (1980).
6. Panayotatos, N. & Wells, R. D. *Nature* **289**, 466-470 (1981).
7. Wu, H.-M. & Crothers, D. M. *Nature* **308**, 509-513 (1984).
8. Mizuuchi, K., Kemper, B., Hays, J. & Weisberg, R. A. *Cell* **29**, 357-365 (1982).
9. Lilley, D. M. J. & Kemper, B. *Cell* **36**, 413-422 (1984).
10. Lafer, E. M., Moller, A., Nordheim, A., Stollar, B. D. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3546-3550 (1981).
11. Potter, H. & Dressler, D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3698-3702 (1978).
12. Bell, L. & Byers, B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3445-3449 (1979).
13. Kallenbach, N. R., Ma, R.-I. & Seeman, N. C. *Nature* **305**, 829-831 (1983).
14. Lilley, D. M. J. & Markham, A. F. *EMBO J.* **2**, 527-533 (1983).
15. Messing, J. & Vieira, J. *Gene* **19**, 269-276 (1982).
16. Frederick, C. A. *et al. Nature* **309**, 327-331 (1984).
17. Crick, F. H. C. & Klug, A. *Nature* **255**, 530-533 (1975).
18. Sobell, H. M., Tsai, C., Gilbert, S. G., Jain, S. C. & Sakore, T. D. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3068-3072 (1976).
19. Levitt, M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 640-644 (1978).
20. Sussman, J. L. & Trifonov, E. N. *Proc. natn. Acad. Sci. U.S.A.* **75**, 103-107 (1978).
21. Twigg, A. J. & Sherratt, D. *Nature* **283**, 216-218 (1980).

Three-dimensional structure of an antigen-antibody complex at 6 Å resolution

A. G. Amit, R. A. Mariuzza, S. E. V. Phillips & R. J. Poljak

D  partement d'Immunologie, Institut Pasteur, 75724 Paris Cedex 15, France

Present understanding of the three-dimensional structure of antibody combining sites is based on X-ray diffraction studies of myeloma immunoglobulins (reviewed in refs 1, 2). The structures of the antigen-binding fragment (Fab) complexes of two of these immunoglobulins with small ligands have also been determined^{3,4}. However, there is no crystallographic information concerning the interactions of an antibody with an antigen, nor do we know the precise structure of antigenic determinants on protein molecules (reviewed in ref. 5). We now report the first structure determination of an antigen-antibody complex at 6 Å resolution. The structure of the complex between hen egg-white lysozyme and the Fab of a

monoclonal anti-lysozyme antibody (D1.3) shows that the combining site of antibodies is not merely a cleft delineated by the complementarity-determining regions of the variable regions of the light and heavy chains, but is a larger area extending beyond it. A correspondingly large area of the antigen makes close contacts with the antibody, in agreement with the notion of a 'topographical' rather than 'sequential' antigenic determinant. The structural basis of cross-reactivities of an antibody with heterologous antigens and the effect of a single amino acid substitution on antigenic specificity can thus be visualized in the structural model presented here.

Hen egg-white lysozyme was chosen as the protein antigen because its three-dimensional structure has been analysed extensively⁶ and because it has served as a model in many immunological studies on the identification of antigenic sites on a protein molecule⁷⁻¹². The production of hybrid cell lines secreting anti-lysozyme antibodies, the purification of the immune complex between Fab D1.3 and hen lysozyme and the crystallization and preliminary crystallographic study of this Fab-antigen complex have been described elsewhere¹³. Crystals, grown from 15–20% polyethylene glycol 8000 solutions, are monoclinic, space group $P2_1$, with $a = 55.6$ Å, $b = 143.4$ Å, $c = 49.1$ Å, $\beta = 120.5^\circ$ and one molecule of complex per asymmetric unit. The diffraction pattern extends to 2 Å resolution using synchrotron radiation.

Three heavy-atom isomorphous substitutions were obtained with K_2PtCl_4 , $K_3F_7UO_2$ and p -OH-mercuribenzenesulphonate. X-ray intensities were measured using a four-circle automatic diffractometer. The conventional R factors after refinement, using a full-matrix least-squares method¹⁴, were 49, 47 and 44%, respectively (details will be published elsewhere). The average figure of merit for the phases of the 1,659 reflections contained in the 6-Å sphere was 0.75. A balsa wood model based on the electron density map calculated with these reflections showed the typical domain structure of Fab¹⁵ and an additional globular region, corresponding to lysozyme, in close contact with Fab.

The map was displayed subsequently on an Evans and Sutherland PS 300 interactive graphics system using the program FRODO (ref. 16), together with the α -carbon backbones of the previously determined hen lysozyme⁶ and Fab New¹⁷ structures. The helical segments of lysozyme clearly corresponded to rods of electron density. The model was fitted to the map using only rigid body rotations and translations (Fig. 1). The lysozyme chain crosses regions of lower density at positions 52–53, 58–59 and at the turns 21–22 and 115–116, but the interpretation is unequivocal. No gross conformational change is evident at this resolution, but the two C-terminal residues seem to have moved and a rearrangement may have occurred at the 21–22 turn. Both of these sites are involved in contacts with the antibody. The variable domains of Fab could be fitted well and the heavy (H) and light (L) chains distinguished by comparison with the Fab New model^{15,17}. In addition, the alternative fitting of the light and heavy chains was not as satisfactory. The constant domain ($C_H1 + C_L$) could be fitted only by rotating it about an axis through residues 103L and 117H at the flexible switch regions, changing the angle between the constant and variable domains from 137° in Fab New to $\sim 180^\circ$, closer to the 166° angle observed in Fab Kol (ref. 18).

The interactions between Fab and lysozyme extend over a

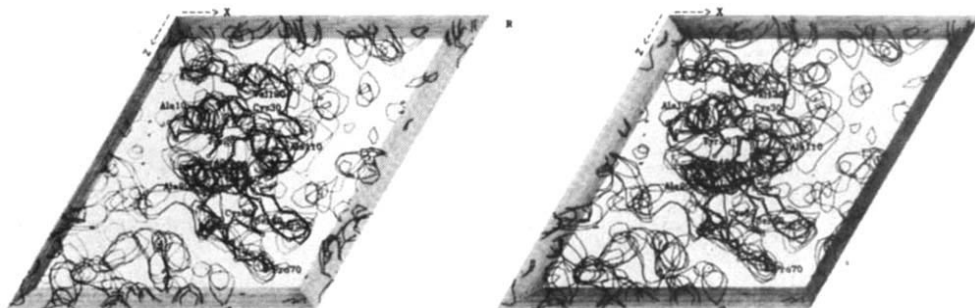
large area (Fig. 2) and involve all six complementarity-determining regions (CDRs)¹⁹ of the antibody. The most obvious contacts in lysozyme occur at positions 19–27 and 116–129, widely separated in its sequence but forming a contiguous patch of about 20×25 Å on its surface. These residues do not correspond strictly to any of the four antigenic sites proposed from studies using polyclonal antibodies^{7,8}. Although side chains are not visible at this resolution, the model (Fig. 2b) particularly suggests contacts to Asn 19, Gly 22, Ser 24, Gly 117, Asp 119, Gln 121 and Arg 125. Possible contacts to side chains such as those of Lys 13, Arg 14 and Arg 21 remain to be verified by a high-resolution analysis, presently under way. Biochemical experiments described below suggest that Gln 121 has an important role in antigen-antibody complex formation. Gln 121 occupies a central position in the region of contact (Fig. 2), in close proximity to CDR1 and CDR3 of the light chain and to CDR3 of the heavy chain, its side chain penetrating the binding cleft of Fab D1.3.

To define the amino acid residues of lysozyme recognized by monoclonal antibody D1.3, we analysed the enzymatic activities of different avian lysozymes in the presence of Fab D1.3 (Fig. 3). Chromatofocusing¹³ on HPLC established that those lysozymes that were inhibited in the assay formed complexes with Fab D1.3. Consequently, we take the enzymatic assay of Fig. 3 as an indicator of complex formation. Whether inhibition of activity is the result of steric hindrance or of an antibody-induced conformational change in the enzyme remains to be determined. Fab D1.3 inhibits the lytic activity of bobwhite quail lysozyme and of hen lysozyme but has little effect on that of California quail lysozyme. The quail lysozymes differ in their amino acid sequences only at positions 68 and 121 (ref. 20). However, there is no apparent difference between the inactivation of hen lysozyme and bobwhite quail lysozyme, which differ at position 68 but share a Gln at position 121, indicating that the replacement of Gln 121 by His in California quail lysozyme interferes with the formation of an antigen-antibody complex. Partridge²¹ and turkey²² lysozymes with a His at position 121 are also not inactivated by Fab D1.3 (data not shown), nor is Japanese quail lysozyme²³ with Asn at position 121, although this latter lysozyme has additional sequence differences at positions 19 and 21, which make close contacts with Fab. These observations are explained readily in terms of the three-dimensional structure presented here, in which Gln 121 occupies a central position in the antigen-antibody contact area. Monoclonal antibody D1.3 therefore recognizes amino acid replacements at position 121, as did polyclonal antisera in a microcomplement fixation assay^{20,21}. Thus, the effect of single amino acid substitutions on antigenic specificity and antigen recognition by the antibody are explained readily in terms of our three-dimensional model.

This work was supported by grants from the CNRS, INSERM and Institut Pasteur. We thank Drs E. M. Prager and A. C. Wilson for lysozyme samples of bobwhite and California quails, Drs J. L. Guénet and A. Perramon for eggs of different bird species and the MRC Laboratory of Molecular Biology, Cambridge, for the use of a computer graphics system.

Note added in proof: Preliminary examination of a 2.8 Å-resolution electron density map confirms the overall assignment of

Fig. 1 Stereo view of the electron density map in a slab 30 Å thick showing the lysozyme molecule. A single contour at $0.5 \text{ e } \text{\AA}^{-3}$ is shown, with the α -carbon skeleton of lysozyme⁶ superimposed. Electron density near the C-terminus at the top corresponds to the combining site of the bound Fab, and that at the bottom to the Fab of a neighbouring complex in the crystal. Most of the polypeptide chain, including all the helical regions, lies inside strong electron-dense features.



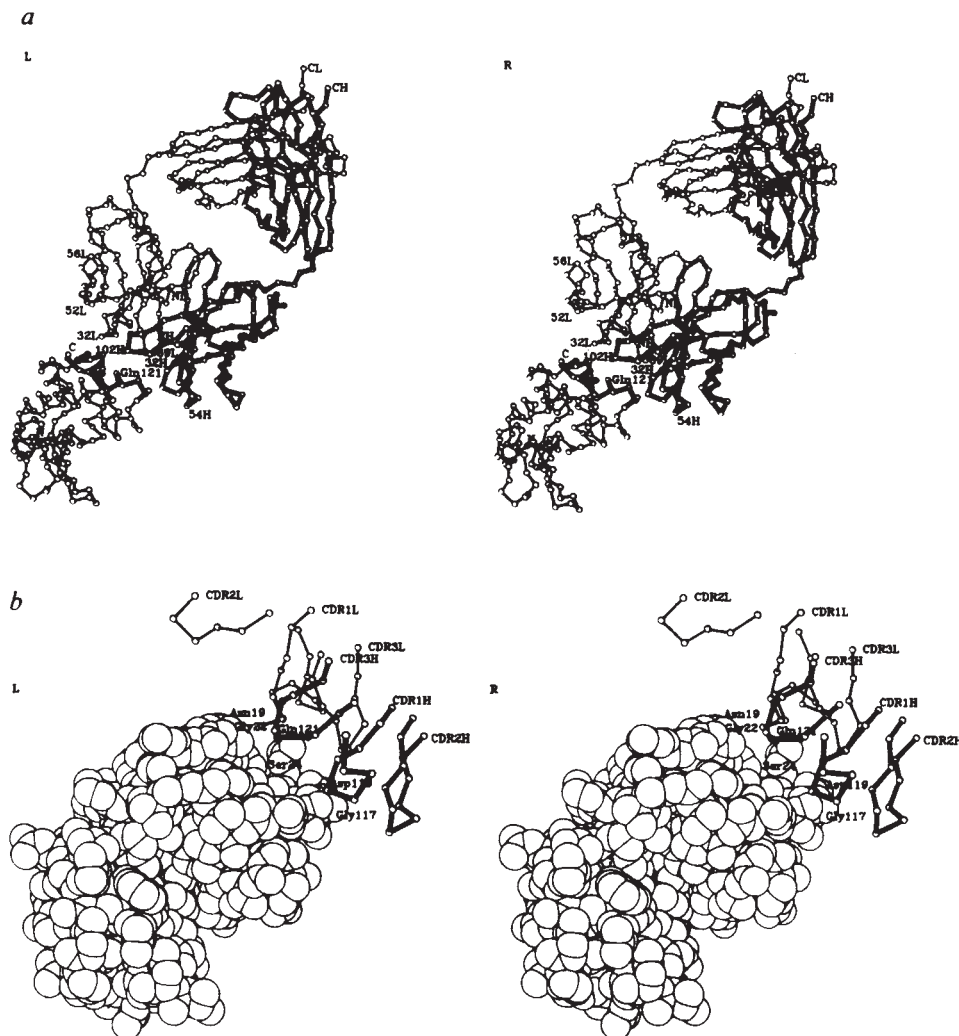


Fig. 2 Preliminary three-dimensional models of the antibody-lysozyme complex. *a*, Stereo diagram of the α -carbon skeleton obtained by fitting the structures of hen lysozyme⁶ and Fab New¹⁷ to the electron density map. *b*, Space-filling representation of lysozyme with all atoms shown as spheres of radius 1.7 Å. Note that in *a*, the establishment of the precise conformation of the CDRs, in particular, that of CDR2 of the light-chain variable region, which is affected by deletion of amino acids 56–63 in Fab New, must await the determination of the amino acid sequences of D1.3 and the calculation of a high-resolution electron density map. In this drawing, CDR2 of the Fab New light chain appears somewhat distant from the antigen. The electron density map, however, indicates that the corresponding CDR of D1.3 makes contacts with lysozyme. The model shown in *b* is a simplification, as it assumes that the orientation of side chains is the same as in crystalline hen lysozyme⁶. The orientation is slightly tilted compared with that of *a* to show the contact region more clearly. Residues most likely to be involved in the interaction with Fab are labelled, except Arg 125, which is hidden, and Lys 116, which lies just in front of Gly 117; the CDR loops of Fab are shown by their α -carbon skeletons. The side chain of Gln 121 lodges between the CDR1 and CDR3 of the light chain and the CDR3 of the heavy chain. The lysozyme active site is the left to the lower right.

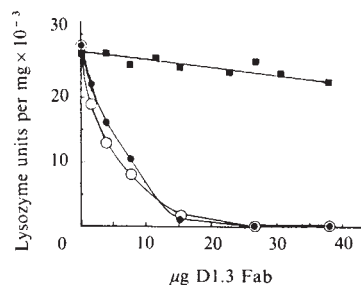


Fig. 3 Inactivation of three lysozymes by D1.3 Fab. The lysozymes of hen (1.3 µg, ○), bobwhite quail (1.2 µg, ●) and California quail (1.0 µg, ■) were incubated in 0.1 M K-phosphate, pH 7.0, with increasing amounts of the Fab fragment of antibody D1.3, at 37 °C for 90 min, in a total volume of 100 µl. Their lytic activity was assayed with *Micrococcus lysodeikticus* cell walls (Sigma; 0.3 mg ml⁻¹ in 0.1 M K-phosphate, pH 7.0) as substrate. A unit of specific enzymatic activity (ordinate) is defined as a decrease in transmittance at 450 nm of 0.001 per min per mg of lysozyme at 22 °C. Incubation with an anti-arsenate Fab, R22.4 (ref. 24), as control, provided the 0 point value, which was 80–95% of that of the activity of untreated lysozymes.

lysozyme residues in contact with the antibody. Recent reports show the importance of segmental mobility in antigenic recognition^{25,26}; the residues of lysozyme recognized by antibody D1.3 are not in regions of above-average mobility.

Received 8 August; accepted 14 October 1984.

- Amzel, L. M. & Poljak, R. J. *A. Rev. Biochem.* **48**, 961–997 (1979).
- Davies, D. R. & Metzger, H. A. *Rev. Immun.* **1**, 87–117 (1983).
- Amzel, L. M., Poljak, R. J., Saul, F., Varga, J. M. & Richards, F. F. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1427–1430 (1974).
- Segal, D. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **71**, 4298–4302 (1974).
- Benjamin, D. C. *et al. A. Rev. Immun.* **2**, 67–101 (1984).
- Blake, C. C. F. *et al. Nature* **206**, 757–761 (1965).
- Arnon, R. in *Immunochemistry of Enzymes and Their Antibodies* (ed. Salton, M. R. J.) 1–28 (Wiley, New York, 1977).
- Atassi, M. Z. *Immunochemistry* **15**, 909–936 (1978).

- White, T. J., Ibrahim, I. M. & Wilson, A. C. *Nature* **274**, 92–94 (1978).
- Kobayashi, T. *et al. Molec. Immun.* **19**, 619–630 (1982).
- Smith-Gill, S. J., Lavoie, T. B. & Mainhart, C. R. *J. Immun.* **133**, 384–393 (1984).
- Mezger, D. W., Ch'ng, L. K., Miller, A. & Sercarz, E. E. *Eur. J. Immun.* **14**, 87–93 (1984).
- Mariuzza, R. A. *et al. J. molec. Biol.* **170**, 1055–1058 (1983).
- Bricogne, G. in *Computational Crystallography* (ed. Sayre, D.) 223–230 (Oxford University Press, 1982).
- Poljak, R. J., Amzel, L. M., Avey, H. P., Becka, L. N. & Nisonoff, A. *Nature new Biol.* **235**, 137–140 (1972).
- Jones, T. A. *J. appl. Crystallogr.* **11**, 268–272 (1978).
- Poljak, R. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **70**, 3305–3310 (1973).
- Matsushima, M. *et al. Molec. Biol.* **121**, 441–459 (1978).
- Kabat, E. A. & Wu, T. T. *Ann. N.Y. Acad. Sci.* **190**, 382–392 (1971).
- Ibrahim, I. M., Prager, E. M., White, T. J. & Wilson, A. C. *Biochemistry* **18**, 2736–2743 (1979).
- Ibrahim, I. M., Eder, J., Prager, E. M., Wilson, A. C. & Arnon, R. *Molec. Immun.* **17**, 37–46 (1980).
- LaRue, J. N. & Speck, J. C. Jr *J. biol. Chem.* **245**, 1985–1991 (1970).
- Kaneda, M., Kato, I., Tominaga, N., Titani, K. & Narita, K. *J. Biochem., Tokyo* **66**, 747–749 (1969).
- Lamoyi, E., Estess, P., Capra, J. D. & Nisonoff, A. *J. Immun.* **124**, 2834–2840 (1980).
- Westhof, E. *et al. Nature* **311**, 123–126 (1984).
- Tainer, J. A. *et al. Nature* **312**, 127–134 (1984).

Terrestrial mass extinctions and galactic plane crossings

IN their attempt to build a case for a "galactic plane crossing" explanation for terrestrial extinction, Rampino and Stothers¹ present a statistical argument that is seriously misleading. The crux of the empirical component of their argument is a claimed agreement between two columns of dates (their Table 1):

Table 1 Dates from ref. 1

Mass extinctions (Myr BP)	Galactic plane crossings (Myr BP)
11	0
37	31
66	64
91	100
144	135
176	166
193	197
217	227
245	259

They begin by calculating the correlation coefficient of these two columns to be $r=0.996$. It is a mistake to attribute any meaning to this number: the correlation of any two monotonically increasing sequences is bound to be high. This reflects only the high serial correlation of the separate series, and in no way indicates any connection between the two series other than a common monotonicity. For example, the correlation between the list of mass extinction dates and the first nine prime numbers (2, 3, 5, 7, 11, 13, 17, 19, 23) is $r=0.986$.

Rampino and Stothers then go on to compare the gaps in the two columns by Student's t -test for matched pairs. There are two serious difficulties in this; the point will be clearer as we introduce some notation. Let $X_1=11, \dots, X_9=245$ be the entries in the first column and $Y_1=0, \dots, Y_9=259$ be the entries in the second column. Then Rampino and Stothers apply Student's t -test to the eight differences $Z_i=(Y_{i+1}-Y_i)-(X_{i+1}-X_i)$. The test statistic is then the average of the Z s divided by what would be an appropriate estimate of standard error of that average if the Z s were independent. But the average of the Z s reduces algebraically to $\bar{Z}=[(Y_9-Y_1)-(X_9-X_1)]/8$; it does not depend on the intermediate values at all. If the two columns spanned the same time interval, it would reduce identically to zero, regardless of the intermediate values. In other words, the Z s are highly correlated because adjacent intervals share common endpoints; the test as performed would be valid only if both series were random walks (hardly a tenable hypothesis, particularly for galactic plane crossings). Consequently, the denom-

inator of the t -statistic grossly overestimates the standard error of the numerator and biases the test statistic towards zero. Even if their test were valid, there is a second difficulty in this aspect of their argument: they interpret a small t -value as evidence in favour of a null hypothesis, when it could as well be held to reflect a paucity of data and a test insufficiently powerful to detect a difference.

S. M. STIGLER

Department of Statistics,
University of Chicago,
Chicago, Illinois 60637, USA

I. Rampino, M. R. & Stothers, R. B. *Nature* **308**, 709-712 (1984).

RAMPINO AND STOTHERS REPLY—Stigler regards as misleading the two statistical tests that we¹ made of the correlation between two observed time series, one containing the dates of the last nine crossings of the galactic plane by the Solar System², and the other consisting of the dates of the last nine major mass-extinction episodes on Earth, which we selected from the marine extinctions record published by Raup and Sepkoski³. Both records cover the past 260 Myr.

In our first test, we computed the correlation coefficient between the two series and obtained $r=0.996$. Stigler correctly points out that any two monotonically increasing series have a numerically high correlation. As a random test case, he correlates the first nine prime numbers with our mass-extinctions time series and obtains $r=0.986$. To make a better comparison, however, we have computed 5,000 random time series, each containing nine dates drawn from a fixed time range (0-260 Myr). These series have been correlated with our mass-extinctions time series. We find agreement with Stigler in that a statistically significant percentage of cases (13%) have $r \geq 0.986$.

On the other hand, only an insignificant 0.4% of the cases computed in these Monte Carlo simulations show $r \geq 0.996$. Furthermore, if we correlate our random time series with the galactic time series, the percentage is even lower, 0.2%. Stigler's mistake, therefore, is to regard $r=0.996$ as being close to $r=0.986$ in the case of two monotonically increasing series with nine members each.

With regard to our second test, the matched-pairs t -test applied to the time intervals, Stigler again makes an invalid criticism, because the data for any matched-pairs t -test at all can be recast and analysed in exactly the way he presents, with the same conclusion. Consider, for example, two random samples drawn from independent populations and containing eight elements each: $A_1, \dots, A_8$ and

$B_1, \dots, B_8$. Enumerate a series of nine X s and a series of nine Y s by computing $A_i = X_{i+1} - X_i$ with any X_1 and $B_i = Y_{i+1} - Y_i$ with any Y_1 . The t -test is now straightforwardly applied to the differences $Z_i = B_i - A_i$. Because the test statistic t is proportional to the average of the Z s, we compute $\bar{Z}$ and find after some manipulation

$$\bar{Z} = [(Y_9 - Y_1) - (X_9 - X_1)]/8$$

which is identical to Stigler's result. Thus, $\bar{Z}$ appears not to depend at all on the intermediate values of the X s and Y s, as Stigler noted. Stigler's result, therefore, is seen to be just a mathematical rearrangement of the data. As long as the sample elements, the A s and B s, are independent and approximately normally distributed, the t -test is valid. In our case, the sample elements are the time intervals in the time series, which are the relevant, physically independent units and are affected by random errors arising from many complex physical and accidental factors for both series¹⁻³. The time intervals are, in fact, found to be approximately normally distributed. Thus, these time series can be regarded as random walks. From the point of view of testing the equality of the averages of the time intervals in the two series, it makes no difference in what order the time intervals are taken, although, if pairs are to be matched, the members of pairs must be kept together.

Our use of the matched-pairs t -test appears, therefore, to be entirely valid. This test is usually considered to be a powerful one, and one for which a sample size of eight is not unusual⁴. (Student⁵ himself used sample sizes of as small as two.) Our original result, based on a two-tailed test, was $t=0.91$ with 7 d.f., therefore $P=0.39$. If, however, we do not pair the time intervals but do allow for the unequal variances in the two samples of time intervals and for the small sample sizes⁶, we obtain $t=0.82$ with effectively 6 d.f., and so $P=0.45$. Alternatively, by simply testing the hypothesis that the average of the time intervals in the mass-extinctions series equals 33 Myr (which is the average time interval in the galactic series), we find $t=1.01$ with 7 d.f., so $P=0.35$. Clearly, even large differences in our test assumptions make little difference to the results.

Thus, we cannot reject the null hypothesis that the averages of the time intervals in the two series are equal. On the other hand, the tests cannot tell us how nearly equal the averages may be. A feeling for what can be safely rejected follows from a further consideration: with the same number of time intervals, another mass-extinctions time series might have a considerably different length. To demonstrate the consequences of adopting the original time series given by Raup and Sepkoski³, in which the last nine mass-

extinction episodes occurred at 11.3, 38, 65, 91, 125, 144, 163, 175 and 194 Myr ago, we have made a new correlation with the galactic time series, finding $r = 0.994$ and, for the matched time intervals, $t = 4.94$ with 7 d.f., for which $P < 0.01$. Thus, the two series are strongly correlated, but in this case they have significantly different averages of the time intervals.

Note that the statistical tests in our original paper were not central to the search for periodicities or the physical arguments there. The approximate periodicity that appeared in both of our listed time series was formally detectable because the variance of the distribution of time intervals in each series was sufficiently small.

MICHAEL R. RAMPINO
RICHARD B. STOTHERS

National Aeronautics and
Space Administration,
Goddard Space Flight Center,
Institute for Space Studies,
2880 Broadway,
New York,
New York 10025, USA

1. Rampino, M. R. & Stothers, R. B. *Nature* **308**, 709–712 (1984).
2. Innanen, K. A., Patrick, A. T. & Duley, W. W. *Astrophys. Space Sci.* **57**, 511–515 (1978).
3. Raup, D. M. & Sepkoski, J. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 801–805 (1984).
4. Klugh, H. E. *Statistics, the Essentials for Research* Ch. 10 (Wiley, New York, 1970).
5. Student [Gossett, W. S.] *Biometrika* **6**, 1–25 (1908).
6. Bailey, D. E. *Probability and Statistics* Ch. 17 (Wiley, New York, 1971).

Hund's rule

BOYD¹ recently suggested an explanation of Hund's rule in terms of electron correlation. A previous investigation of this approach² gave rise to a richer interpretation of two-electron spectra in terms of an alternating rule, in place of the usual Hund rule first noted by Russell and Meggers³ for the spectrum of scandium. The alternating rule states that, within a given configuration, the energy ordering of singlet and triplet states reverses each time the angular momentum changes by one unit. For the term of greatest angular momentum, the triplet lies lowest. The rule reliably orders singlet–triplet energy levels in ~90% of known cases.

The alternating rule follows from the assumption that the state with the larger average value of the inter-electronic angle, θ_{12} , has less shielding and therefore a lower energy. This is not far from Boyd's approach. One might expect a triplet always to have less shielding, owing to the Fermi hole, but wavefunction antisymmetry can give rise to features in the singlet which resemble the triplet hole². As the charge density $\rho(\theta_{12})$ has been found to obey the alternating rule², the Fermi hole argument is too simplistic.

The alternating rule helps in explaining cases deviating from Hund's first rule, but

many exceptions are known. When the second rule fails, a singlet can be the lowest term in a configuration. An example is the $2p3d$ configuration of carbon, where the 1D lies below the 3F , contrary to Hund's first rule. The alternating rule, however, is obeyed: the 1D lies below the 3D , and the 3F is below the 1F . For the $npn'd$ configuration, Hund's rules work only in ~50% of the cases² and many of the exceptions may be caused by a failure of Hund's second rule.

Recent studies of electron distributions in two-electron atoms^{4,5} are beginning to yield a new and more comprehensive picture. Some states, for example, the doubly-excited states of helium-like atoms, exhibit very strong angular correlations, much greater than in singly-excited states, to the extent that collective, quasi-molecular quantization, with near-rigid rotations and bending and stretching vibrations, describes the electron distribution. Even the alkaline-earth atoms and alkali negative ions show considerable A–B–A molecular behaviour, although these systems have less rigid structures than He⁺⁺.

That radial correlation sometimes dominates the energy orderings is difficult to reconcile with shielding concepts². The $2s3s$ states of helium have similar angular electron distributions, much like that of the $2s^2$, 1S state, but the radial distributions are all quite different. Both the $2s3s$ states have strong angular correlation expected for a linear quasi-molecular system, with $\rho(\theta_{12})$ centred around π . The triplet corresponds to one quantum of excitation in the antisymmetric stretching mode for the A–B–A molecule; it has a lower energy than the singlet because the antisymmetric stretching has a lower frequency when A is lighter than B. The $2s2p$, 3P helium state, with a strong maximum in $\rho(\theta_{12})$ near π , corresponds to the first-excited rotor state, with no bending excitation. The angular distribution in the $2s2p$, 1P state is very different, with the maximum at slightly more than $\pi/2$, and with finite probability density at $\theta_{12} = 0$. This singlet is one of two partners in a nearly degenerate pair with one quantum in the bending vibration; the other is the lower energy $2p^2$, 3P , having almost the same angular distribution, but no electron density at $\theta_{12} = 0$.

In the quasi-molecular picture, the appropriate comparisons, as exemplified above, may be between states of different configurations. When electron correlation and configuration interaction are large, it is inappropriate to classify states in terms of a one-electron configuration, and Hund's rule loses its meaning. Nevertheless, it is interesting to ask under what circumstances a shielding argument can predict the energy orderings. The alternating rule has many exceptions for the $npn'd$, P and D terms, and future considerations of this problem should perhaps focus on physical interactions, such as collective motion kinematics and shielding, rather than on quantum numbers associ-

ated with a specific model and representation.

J. W. WARNER
R. STEPHEN BERRY

Department of Chemistry,
The University of Chicago,
Chicago, Illinois 60637, USA

1. Boyd, R. J. *Nature* **310**, 480–481 (1984).
2. Warner, J. W., Bartell, L. S. & Blinder, S. M. *Int. J. Quantum Chem.* **18**, 921–932 (1980).
3. Russell, N. H. & Meggers, W. F. *Sci. Pap. natn. Bur. Stand.* **22**, 364–371 (1927).
4. Rehms, P. D. & Berry, R. S. *Chem. Phys.* **38**, 247–275 (1979).
5. Ezra, G. & Berry, R. S. *Phys. Rev. A* **28**, 1974–1988 (1983).

Frost rings in trees and volcanic eruptions

LAMARCHE and Hirschboeck¹ find good agreement between the timing of notable frost events in the western United States and the occurrence of major volcanic eruptions worldwide from AD 1500 onwards. Notable frost events are defined as those damaging trees (seen in tree rings) at two or more localities or in 50% or more of sampled trees in any one locality. For the periods 1882–1968 and 1500–1880 they find, by considering coincidences of these phenomena in 3-yr segments, χ^2 -values of 16.0 and 14.7, respectively, in frost versus volcano contingency tables, indicating significance at the 99.95% confidence level.

I have the following comments to make: a recomputation of χ^2 is possible using only the occasions (underlined in Table 1 of ref. 1) of geographically-widespread frost damage (that is, in both the Great Basin and the Rocky Mountains regions, $>10^\circ$ longitude apart). The results are: 1881 onwards, $\chi^2 = 6.7$, just significant at 99% confidence level; 1500–1880, $\chi^2 = 7.8$, significant at 99% confidence level.

The authors' selection of volcanic events, after Lamb², appears to be inconsistent at a few points. In general, they selected events printed² in bold type, indicating a consensus among compilers as to the importance of the events, with $DVI/E_{\max} \geq 1,000$ for the eruption or combined eruptions. Events, with the latter parameter in bold type and $\geq 1,000$ but with the name of the event in normal type (for example, Pogrumni, 1795), were omitted. To have followed this objective procedure consistently would have been acceptable but the following deviations were noted in particular: (1) Awu and Gunung (both 1641, combined $DVI/E_{\max} = 1,500$) were omitted. (2) The 1785 eruption of Vesuvius was not assigned a DVI by Lamb and should have been omitted. (3) Hekla (1845) and Amargura (1846) (combined $DVI/E_{\max} \approx 1,800$) were omitted. (4) A group of four eruptions (1886–88) (combined $DVI/E_{\max} = 1,100$ (estimated)² in 1888–90) was omitted. Item (2) will not affect the authors' results because Vesuvius was grouped with other eruptions, themselves with $DVI/E_{\max} > 1,000$. However, according to Table 1 in ref. 1, the events in (1), (3) and

(4) were not followed by frosts and so χ^2 will be reduced if they are included.

The eruptions of 1752 and 1755, both with $DVI/E_{\max} > 1,000$, should have been in separate triads. Neither triad experienced frost, so χ^2 will be reduced further.

Inclusion of the volcanic events listed in (1) and (3) and separation of the 1752 and 1755 eruptions reduces the χ^2 for 1500–1800, computed using all frost events in Table 1, ref. 1 from 14.7 to 10.9; the significance level falls from 99.95 to 99.9%. Similarly, inclusion of (4) reduces the all-frost-events χ^2 for 1882–1968 to 11.9 and its significance to 99.9%.

The severest test is to treat the volcanic events as described in the previous paragraph and to recompute χ^2 using only the occasions of widespread frost. The results are: 1500–1880, $\chi^2 = 6.2$ (98% level); 1882–1968, $\chi^2 = 5.0$ (95% level).

For the 1882–1968 period, widespread frosts followed only Katmai and Agung (see ref. 1, Table 1) in which $DVI/E_{\max} < 1,000$. A procedure more consistent with 1500–1880 would have excluded these eruptions. A repeat of our severest test would then give null results: no widespread frost followed the major eruptions ($DVI/E_{\max} \geq 1,000$) in 1883, 1886–88 or 1902.

I conclude that a careful examination of Lamb's² volcanic record shows that the connection with frost damage to trees in the western United States is not as significant as LaMarche and Hirschboeck¹ indicate. This is particularly evident if occasions of widespread frost only are considered. However, it cannot be denied that some connection is still apparent for the period 1500–1880.

D. E. PARKER

*Meteorological Office,
London Road, Bracknell,
Berkshire RG12 2SZ, UK*

1. LaMarche, V. C. Jr. & Hirschboeck, K. K. *Nature* **307**, 121–126 (1984).

2. Lamb, H. H. *Phil. Trans. R. Soc. A* **266**, 425–533 (1970).

LAMARCHE AND HIRSCHBOECK REPLY—We thank Parker for his comments on our article but can agree with only some of his points.

There were indeed some minor inconsistencies in our choice of major volcanic eruptions for comparison with notable frost-ring events in the western United States. In particular, Awu and Gunung (1641), Hekla and Amargura (1845–46) and a group of four eruptions in 1886–88 ought to have been included in our tabulation and were omitted by mistake. As Parker points out, recomputation of χ^2 , after adding these data, reduces the significance of coincidences of eruptions and frost rings. However, they remain highly significant at the 99.9% confidence level for both the 1500–1880 and 1882–1968 sub-periods.

We believe that his recomputation of χ^2 , using a small subset of frost-damage

dates, may be premature. The unstated assumption here is that widespread meteorological events are more likely to be linked physically to the influence of a hemispheric-scale volcanogenic-aerosol veil in the stratosphere. This may be true but, as documented more fully in references cited in our article, the timing of the frost in relation to the progress of frost-hardening in the plant tissues may also be critically important. If unseasonable cold occurs too late in the growing period, even by a few days, no permanent damage may result. Therefore, the record of a major meteorological event may be restricted geographically, or even totally absent, depending on its timing and on antecedent climatic conditions affecting the hardening of plant tissues.

Finally, the exclusion of the Katmai (1912) and Agung (1963) events in a recalculated χ^2 analysis is not justified. As we point out in a footnote to our Table 1, estimated aerosol injections for these eruptions have been revised upward since Lamb published his tabulations.

VALMORE C. LAMARCHE JR*
KATHERINE K. HIRSCHBOECK†

* *Laboratory of Tree-Ring Research,
The University of Arizona, Tucson,
Arizona 85721, USA*

† *Department of Geography and
Anthropology,
Louisiana State University,
Baton Rouge, Louisiana 70803, USA*

Mediation of slow-inhibitory postsynaptic potentials

It has been claimed by Cole and Shinnick-Gallagher¹ that the slow-inhibitory postsynaptic potential (s-i.p.s.p.) in a mammalian sympathetic ganglion is due to the monosynaptic activation of muscarinic receptors. This conflicts with evidence for a disynaptic mechanism in which acetylcholine (ACh) acts muscarinically on intraganglionic interneurons (SIF cells) to release the direct inhibitory transmitter, dopamine².

The three lines of crucial evidence needed to establish the identity of a synaptic transmitter³ had been developed for dopamine: (1) Dopamine and its synthesizing enzymes are present in appropriate presynaptic structures^{4,5}. (2) Postsynaptic actions by catecholamines including dopamine are appropriate to those in an s-i.p.s.p., that is, hyperpolarization without appreciable change in membrane conductance (g_m)^{2,6–8}. (3) Dopamine is depleted, presumably released, from SIF cells and their terminations by preganglionic nerve impulses or by muscarinic agonists⁵. Furthermore, depletion of dopamine is accompanied by a loss of s-i.p.s.p., while subsequent application of dopamine leads to its re-uptake by SIF cells and restores s-i.p.s.p. response^{2,5,9}. This specific restoration of

s-i.p.s.p. by applied dopamine is particularly difficult to explain on a monosynaptic, direct muscarinic hypothesis.

The requirement of a second transmitter, whether dopamine or some other compound, has been established by the effects of a low Ca^{2+} (0.1–0.2 mM)/high Mg^{2+} (6 mM) medium; the abolition of presynaptic release of transmitters is accompanied by elimination of the muscarinic hyperpolarizing (but not slow depolarizing) response to ACh or its agonists^{2,10,11}. Cole and Shinnick-Gallagher question this conclusion because, unlike others^{2,10,11}, they find neurally-evoked s-i.p.s.ps are eliminated much faster than ACh hyperpolarization. Such a temporal discrepancy need not exclude a required release of a second transmitter by the ACh; it could simply reflect a faster reduction of Ca^{2+} at preganglionic terminals on the putative interneurons than at interneuronal terminations on ganglion cells, especially with the more drastic conditions of 0 Ca^{2+} with EGTA added¹. Similarly, the reported inability of tetrodotoxin (TTX) to abolish ACh hyperpolarization¹ does not exclude a required second transmitter. TTX blocks Na^+ action potentials, not synaptic release. As previously suggested¹², extra-synaptic muscarinic receptors may be located on the interneuronal terminations near their release sites at ganglion cells, or the fine efferent fibres of the interneurons (SIF cells) might be conducting 'calcium spikes' that do not block with TTX. Furthermore, a suppressing effect of TTX on ACh hyperpolarization, opposite to that of Cole and Shinnick-Gallagher, was previously reported in similar experimental conditions¹¹.

The issue of the electrogenic mechanism of the s-i.p.s.p. is a separate one. First, Cole and Shinnick-Gallagher did not test whether their observed small increase in g_m during an s-i.p.s.p. would also accompany dopamine-induced hyperpolarization in their experimental conditions; therefore, this evidence does not argue against dopamine mediation. Second, the apparent absence of an increase in g_m when resting membrane potentials fall below (less negative than) –60 mV (Fig. 3b, in ref. 1) conflicts with other findings^{13–15} and suggests that experimental conditions were significantly special. Third, there might be two subsets of receptors of s-i.p.s.p. hyperpolarization, with only one involving an increase in g_K by a direct ACh action that is functionally evident in certain test conditions⁸.

Finally, there remains some ambiguous pharmacological evidence. The ability of a low concentration (1 μ M) of the adrenergic antagonist phentolamine to block hyperpolarizing responses to exogenous catecholamines but not the neurally induced s-i.p.s.p. is taken to be evidence against mediation by a catecholamine transmitter. But competitive adrenergic antagonists at higher concentrations have in fact been shown to

depress both muscarinic hyperpolarization² and s-i.p.s.p.s^{9,16}. Indeed, it is not uncommon to find large discrepancies in required concentrations of antagonists tested against applied compared with neurally-released transmitters. As noted¹⁶, this can be explained as due to much greater peak concentrations of transmitter when synaptically released at the receptor sites.

Thus, the more crucial positive evidence for mediation of s-i.p.s.p. by a second transmitter, dopamine, especially in normal intact mammalian sympathetic ganglia (reviewed in refs 2, 10, 12, 16, 17) cannot be ignored and must be dealt with by any alternative hypothesis. On the other hand, the ambiguous pharmacological evidence^{1,2,10,16} can be explained on other grounds, and the significance, for transmitter identity, of the reported small increase in g_m (ref. 1) remains to be clarified.

BENJAMIN LIBET

Department of Physiology,
University of California,
San Francisco, California 94143, USA

1. Cole, A. E. & Shinnick-Gallagher, P. *Nature* **307**, 270-271 (1984).
2. Libet, B. *Fedn Proc.* **29**, 1945-1956 (1970).
3. Werman, R. *Comp. Biochem. Physiol.* **18**, 745-766 (1966).
4. Björklund, A., Cegrell, L., Falck, B., Ritzén, M. & Rosengren, E. *Acta physiol. scand.* **78**, 334-338 (1970).
5. Libet, B. & Owman, Ch. *J. Physiol., Lond.* **237**, 635-662 (1974).
6. Kobayashi, H. & Libet, B. *J. Physiol., Lond.* **208**, 353-372 (1970).
7. Dun, N. & Nishi, S. *J. Physiol., Lond.* **239**, 155-164 (1974).
8. Ivanov, A. Ya. & Skok, V. I. *J. Auton. nerv. Syst.* **1**, 255-263 (1980).
9. Libet, B. & Tosaka, T. *Proc. natn. Acad. Sci. U.S.A.* **67**, 667-673 (1970).
10. Nishi, S. in *Integrative Functions of the Autonomic Nervous System* (eds Brooks, C. McC., Koizumi, K. & Sato, A.) 223-233 (University of Tokyo Press & Elsevier, Amsterdam, 1979).
11. Dun, N. J. & Karczmar, A. G. *J. Pharmac. exp. Ther.* **215**, 455-460 (1980).
12. Libet, B. in *SIF Cells: Structure and Function of the Small Intensely Fluorescent Sympathetic Cells* (ed. Eränkö, O.) 163-177 (U.S. Government Printing Office, Washington DC, 1976).
13. Hashiguchi, T., Ushiyama, N., Kobayashi, H. & Libet, B. *Nature* **271**, 267-268 (1978).
14. Hashiguchi, T., Kobayashi, H., Tosaka, T. & Libet, B. *Brain Res.* **242**, 378-383 (1982).
15. Constanti, A. & Brown, D. A. *Neurosci. Lett.* **24**, 289-294 (1981).
16. Ashe, J. H. & Libet, B. *Brain Res.* **242**, 345-349 (1982).
17. Kuba, K. & Koketsu, K. *Prog. Neurobiol.* **11**, 77-169 (1978).

SHINNICK-GALLAGHER AND COLE
REPLY—In our report¹, we provided evidence suggesting that the slow-inhibitory postsynaptic potential (s-i.p.s.p.) in mammalian sympathetic ganglia is due to the monosynaptic activation of muscarinic receptors. Our data did not support Libet's hypothesis of a disynaptic process for the s-i.p.s.p. resulting from muscarinic activation of an interneurone (SIF cell) which releases dopamine².

Here we analyse whether acetylcholine (ACh) or dopamine fulfill the six criteria for establishing the identity of the neurotransmitter³ mediating the s-i.p.s.p.: (1) Localization: ACh is present preganglionically⁴ and dopamine fluorescence is observed in SIF cells⁵, but whether the SIF cell is an appropriate anatomical sub-

strate for the s-i.p.s.p. is questionable (see ref. 6). (2) Release: ACh is released during presynaptic stimulation⁴, but release of ³H-dopamine from sympathetic ganglia has not been demonstrated⁶ and loss of the s-i.p.s.p. may occur with no change in dopamine fluorescence⁵. Furthermore, restoration of the s-i.p.s.p. does not depend on extrinsic dopamine⁶. (3), (4) Enzymes: Synthesizing and metabolizing enzymes for ACh are present pre- and postsynaptically⁴, respectively, and some dopamine metabolism occurs within the ganglion⁴, but this does not necessarily imply a correlation with the s-i.p.s.p. (5) Synaptic mimicry: ACh causes membrane hyperpolarization^{1,4,7}, having a reversal potential similar to the s-i.p.s.p.^{1,8}. We observed small increases in membrane conductance (g_m) during these responses using instantaneous (not steady state) voltage deflections reflecting initial segment spikes, not rectification in the depolarizing direction. Similar small increases in g_m have been reported by others^{7,8} and can be detected on close inspection in earlier records^{9,10}. Catecholamines do hyperpolarize, but dopamine was the least effective agonist^{11,12} and its effects were mediated through α_2 -adrenoceptors, not dopaminergic receptors¹²; dopamine seemed an unlikely candidate and was not analysed. (6) Identical pharmacology: Atropine blocks both the s-i.p.s.p. and ACh hyperpolarization^{1,6,12} but there is little evidence¹³ that α -adrenoreceptor antagonists block the s-i.p.s.p. High concentrations of competitive adrenergic antagonists only slightly depress s-i.p.s.p.^{10,13} but do attenuate action potentials¹⁴ and other slow potentials^{6,13,14}. Experiments testing whether non-competitive antagonists were more effective at greater concentrations of synaptically released neurotransmitter¹³ are inconclusive, as the antagonists inhibit muscarinically mediated responses¹⁵ and bind to muscarinic receptors and calcium channels but have a greater affinity for the calcium channel¹⁶. We^{1,12} and others^{6,8,14,17} have found that catecholamine antagonists cause no significant depression of the s-i.p.s.p., but these concentrations of antagonists can completely abolish synaptic responses known to be mediated by catecholamines in other neurons¹⁸.

Experiments examining the requirement for a second transmitter are neither a criterion nor definitive, because the treatments interfere with membrane mechanisms. We¹ and others⁸ observed that an ACh hyperpolarization persisted when the synaptic response, the s-i.p.s.p., was blocked by a Ca^{2+} -free, high- Mg^{2+} , EGTA medium; continued superfusion depressed the ACh hyperpolarization. However, in no other previous reports were those experiments performed on the same neurone or in the same preparation. There is no experimental evidence that the temporal discrepancy could be due to

diffusional barriers. On the other hand, muscarinic activation of a Ca^{2+} -dependent K^+ response that persists in Ca^{2+} -free media and is subsequently abolished on repeated application of ACh has been reported in other tissues^{19,20}. We observed only that the ACh hyperpolarization persisted in tetrodotoxin (TTX) at the time the synaptic response, the s-i.p.s.p., was blocked¹. We know of no evidence suggesting that synaptic transmission can occur in the presence of TTX at interneuronal or any known synapse. In another study⁷, the ACh hyperpolarization was present, albeit depressed, when the antidromic action potential was abolished. This suggests that TTX may be affecting the membrane mechanism of ACh by blocking resting Na^+ conductance and indirectly affecting the Na^+ pump, intracellular Ca^{2+} concentration and inward Ca^{2+} currents²¹. Similar effects on ACh hyperpolarizations have been observed when extracellular Na^+ was replaced with Li^+ in amphibian sympathetic ganglia²².

Clearly, the evidence indicates that ACh, not dopamine fulfils all the criteria for the neurotransmitter mediating the s-i.p.s.p. in this^{1,8} and other autonomic ganglia²³⁻²⁵. There is little basis for a disynaptic hypothesis for the s-i.p.s.p. at any autonomic synapse.

PATRICIA SHINNICK-GALLAGHER
Department of Pharmacology,
University of Texas Medical Branch,
Galveston, Texas 77550, USA

ALISON E. COLE
Laboratory of Neurophysiology,
NIH-NINCDS,
9000 Rockville Pike,
Bethesda, Maryland 20205, USA

1. Cole, A. E. & Shinnick-Gallagher, P. *Nature* **307**, 270-271 (1984).
2. Libet, B. *Fedn Proc.* **29**, 1945-1956 (1970).
3. Werman, R. *Comp. Biochem. Physiol.* **18**, 745-766 (1966).
4. Skok, V. I. *The Physiology of Autonomic Ganglia* (Igaku Shoin, Tokyo, 1973).
5. Libet, B. & Owman, Ch. *J. Physiol., Lond.* **237**, 635-662 (1974).
6. Dun, N. J. *Fedn Proc.* **39**, 2982-2989 (1980).
7. Dun, N. J. & Karczmar, A. G. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4029-4032 (1978).
8. Nakamura, T. & Nishi, N. *Neurosci. Lett.* **31**, S78 (1982).
9. Kobayashi, H. & Libet, B. *Proc. natn. Acad. Sci. U.S.A.* **60**, 1304-1311 (1968).
10. Ivanov, A. Ya. & Skok, V. I. *J. autonomic nerv. Syst.* **1**, 255-263 (1980).
11. Brown, D. A. & Caulfield, M. P. *Br. J. Pharmac.* **65**, 435-445 (1979).
12. Cole, A. E. & Shinnick-Gallagher, P. *J. Pharmac. exp. Ther.* **217**, 440-444 (1981).
13. Ashe, J. H. & Libet, B. *Brain Res.* **242**, 345-349 (1982).
14. Dun, N. J. & Karczmar, A. G. *J. Pharmac. exp. Ther.* **215**, 455-460 (1980).
15. Boyd, H. et al. *Br. J. Pharmac.* **20**, 418-435 (1963).
16. Richelson, E. & El-Fakahany, E. *Biochem. Pharmac.* **30**, 2887-2891 (1981).
17. Nakamura, J. *Kurume med. J.* **25**, 241-253 (1974).
18. Egan, T. M., Henderson, G., North, R. A. & Williams, J. T. *J. Physiol., Lond.* **345**, 477-488 (1983).
19. Putney, J. W. Jr *J. Physiol., Lond.* **268**, 139-149 (1977).
20. Trautmann, A. & Marty, A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 611-615 (1984).
21. Standen, W. B. *Nature* **293**, 158-159 (1981).
22. Smith, P. A. *J. Physiol., Lond.* **347**, 377-395 (1984).
23. Horn, J. P. & Dood, J. *Nature* **292**, 625-627 (1981).
24. Hartzell, H. C., Kuffer, S. W., Stickgold, R. & Yoshikami, D. *J. Physiol., Lond.* **271**, 817-846 (1977).
25. Gallagher, J. P., Griffith, W. H. & Shinnick-Gallagher, P. *J. Physiol., Lond.* **332**, 4732 (1982).

Philosophy of mind

The Reith Lectures for 1984 were delivered by Professor John Searle under the title 'Minds, Brains and Science'. Stuart Sutherland discusses some of the issues raised in the broadcasts.

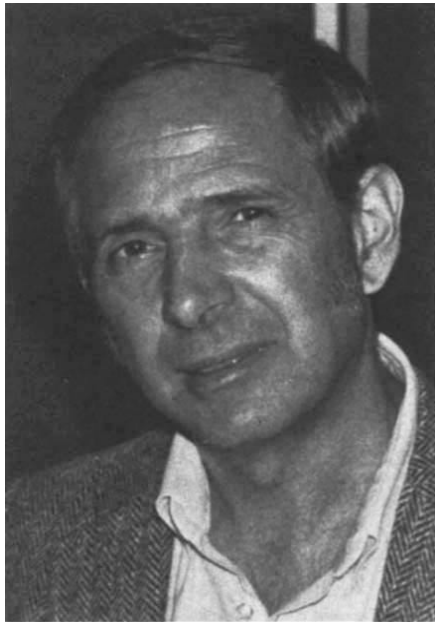
FOR SOME time, one of the main occupations of philosophers has been to tease scientists. Richard Rorty and others have argued that science is valid only in terms of the conventions adopted at a given time, while Paul Feyerabend goes further and believes that within science (and presumably also within philosophy) "anything goes". In the recent series of Reith Lectures, broadcast on BBC Radio 4, Professor John Searle confined his teases to artificial intelligence: as befits a philosopher forged at Oxford and honed at Berkeley, he produced a carefully thought-out view of minds, brains and science. Although provocative, he was for the most part refreshingly clear, so much so that it is possible to refute him on certain points.

Searle believes that mental terms are simply another way of describing the activities of the brain, a view that has the merit of plausibility, though not of originality. In the lectures he likened the emergence of mental phenomena, such as consciousness, intention or desire, to the emergence of such properties as solidity or transparency which are produced by the arrangement of the molecules in a material. But he ignored an important difference: we understand rather fully how certain molecular arrangements can give rise to the higher level attributes of solidity or transparency, but we have no idea whatever of how any arrangement of neurones could give rise to consciousness. (This point was also made by Colin McGinn in an otherwise rather disappointing discussion of the talks during which Searle and his opponents rarely seemed to meet.) One may conclude that Searle has merely by-passed the mind-body problem; he has not made it go away.

Searle attacked three of the claims made by some of the AI fraternity. First, he maintained that no computer system could ever exhibit consciousness, since consciousness is a property only of brains. This claim is reasonable, although Searle defended it by assertion rather than argument.

His second counter-claim was that computers can never exhibit mental activity — they can never have knowledge, even unconscious knowledge. His main argument stemmed from his well-known analogy of the Chinese room. He asked the listener to imagine himself sitting in a room which contains a supply of Chinese symbols and a rule book in English for manipulating them. When the person in the room is handed a string of Chinese symbols, he assembles a sequence of his own symbols in a manner dictated by his

rule book. If the rules were comprehensive enough, and if the strings he received were genuine questions in Chinese, then the strings he assembled could be correct answers. The rules might correspond to those embedded in a computer program that answered questions in Chinese. But clearly the person manipulating the symbols does not understand Chinese nor *a fortiori* does the computer program. To



John Searle — a mind of his own

understand Chinese a person must not merely be able to manipulate the symbols according to rules, he must know the meaning of those symbols. Searle argued that the Chinese room analogy shows that so long as someone or something, whether it be a person or a computer, is automatically following a set of rules (a syntax) while unaware of their meaning (a semantics), he or it cannot be said to have a knowledge of whatever it is the symbols refer to.

The argument is highly ingenious, but it raises two problems. First, if the difference between computer programs and brains is that the latter have semantic knowledge while the former do not, it was surely incumbent on Searle to consider how it is that the brain acquires its semantics; but he did not even mention this question. Second, we would be more tempted to think of computers as having knowledge (a semantics), if they interacted with the world through sensors and mechanical limbs, if they could "learn" from these interactions and if their behaviour was governed by goals in the way that human behaviour is. Searle's reply to this point was too skimmed to carry conviction.

Searle's third and least plausible claim was that not only can computers never have mental activities, they cannot even simulate such activities. Now if the brain operated according to rules, it would clearly be possible to simulate its workings in computer programs. But since the brain and mental activity are one and the same according to Searle, he was driven to the desperate device of denying that the brain operates according to rules: indeed he suggested that no unconscious calculation is undertaken by the brain. But, to take one example, it is known that certain neurones in the striate cortex compute a function on parts of the retinal image. His comment on how the brain does mediate thoughts was distinctly unhelpful — "The brain just does them". The extent to which the brain carries out computations can surely be decided only by investigation not by philosophical *fiat*, and if it does not work like this, it is extremely hard, as Searle's own remarks made clear, to think of any other way in which it might operate.

Searle also attempted to show that there can be no social "science", at least in the meaning of the term arrogated by physical scientists. He claimed that there can be no "systematic correlation between phenomena identified in social and psychological terms and phenomena identified in physical terms". Anything — paper, metal, pigs — can serve as money: it depends only on the attitudes of the community. But this surely does not imply that there are not highly specific and, at least in theory, identifiable processes going on in the brain when someone uses the concept of money. It only means that these processes are likely to be complicated and to be interdependent on many other processes that underlie people's systems of belief. Moreover, there is an implicit contradiction in denying that certain concepts cannot have a systematic representation in the brain while claiming that mental activity *is* the working of the brain.

In his final lecture, which was on freedom of the will, Searle acknowledged defeat. If the mind is merely the brain under another guise, and if the physical world is determined, then there appears to be no room for freedom of the will. Searle said, with commendable honesty, "For reasons that I don't really understand, evolution has given us a form of experience of voluntary action where the experience of freedom is built into the very structure of ... voluntary ... behaviour". It is perhaps curious that he did not begin his lectures by saying the same thing about consciousness over which the same problem arises in an equally acute form. Despite Professor Searle's sleight of hand, the problem of consciousness continues to defeat the human mind. □

Stuart Sutherland is Director of the Centre for Research on Perception and Cognition at the University of Sussex.

Story of a synthetic polypeptide

Jen Tsi Yang

Poly(γ -Benzyl-L-Glutamate) and Other Glutamic Acid Containing Polymers.

By H. Block.

Gordon & Breach: 1984. Pp.215. \$49.

Poly(γ -benzyl-L-glutamate) is unique, the first synthetic polypeptide studied that is a helical rod in helicogenic solvents and a "random coil" in "good" solvents. In the 1950s several laboratories in Britain, the United States and Israel were especially active in investigating this polypeptide. They were heady times. One morning in 1954, I remember, I was with Professor P. Doty at Harvard when a co-worker of his rushed in with a new set of data. Doty paused and plotted them on a graph; the data fitted nicely on a viscosity-molecular weight plot for rigid rods with the dimensions of α -helices — an exciting moment in an exciting era of biophysical chemistry.

Work was soon extended to poly(L-glutamic acid), prepared by removing the protective benzyl group. Since then aspects of the helix-coil transition, both experimental and theoretical, have fascinated many researchers and Professor Block's book covers much of the story to date. However, the water-soluble poly(L-glutamic acid), which is equally versatile in its conformations as poly(γ -benzyl-L-glutamate) and is related to protein conformation, is hardly noted.

Two-fifths of the text is devoted to solution properties (Chapter 5) and about one-third to the solid state and surface properties (Chapter 6), these chapters following a brief introduction and accounts of syntheses and reactions. A short chapter on commercial exploitation completes the book, but it is poly(γ -methyl-L-glutamate) rather than poly(γ -benzyl-L-glutamate) — the cost of which remains an obstacle — that has found use in making fibres, films and leather substitutes. One attractive feature is the bibliography which contains over 900 references. Many of these are now classics, and newcomers to the field will be able to use this part of the book to dig out useful background reading.

Professor Block surveys various topics from organic syntheses to physical studies, but in such a short volume the coverage of each is bound to be somewhat sketchy. Recent developments in the synthesis of sequential polypeptides are listed, for instance, but the brief discussion fails to do them justice. Readers will also need to be familiar with the various techniques to follow the text in detail. Often the terminology is not as well defined as it might have been; for example the narrow Poisson distribution for polydispersity of the amine-initiated polymer formation is not

only poorly defined but the probable causes for departures from it are not considered. The kinetic treatment of copolymerizations is also somewhat scant, and while the formal results of the Zimm-Bragg theory for the helix-coil transition are presented, a simple heuristic quasi-chemical treatment would have made the whole phenomenon more comprehensible. Happily, the subsequent calorimetric studies of this process are taken up in more detail.

The book contains only nine figures, surely too few for this is in part a visual subject. In particular I would like to have seen a Zimm plot for light-scattering data — the classical method of determining the size and shape of the polypeptide as helical rods in helicogenic solvents.

I must also say that I find questionable

the basic rationale of this series of monographs, which "excludes macromolecules of purely academic interest". Before the development of recombinant DNA techniques, and the ensuing rise of a biotechnology industry, such biopolymers as polynucleotides, let alone restriction endonucleases, would have been judged as "useless". That is to say, today's laboratory curiosity may be the gleam in the eye of tomorrow's venture capitalists. One can hope that the fascinating story of poly(γ -benzyl-L-glutamate) and poly(L-glutamic acid) will nonetheless simulate further applied research and that Professor Block's book will help towards that end. □

Jen Tsi Yang is Professor of Biochemistry at the University of California, San Francisco.

Becoming a mammal

Gillian King

The Evolution of Mammalian Characters.

By D.M. Kermack and

K.A. Kermack. *Croom Helm*: 1984. Pp.149. £15.95.

THE mammal-like reptiles, or synapsids, have not had the exposure (popular or academic) with which their bigger and brasher relatives, the dinosaurs, have been favoured. When, for instance, did we last see a voluptuous film-star in fur bikini being chased by *Inostrancevia*? Even the names are not very familiar.

Yet the synapsids have left an excellent fossil record and present us with mysteries every bit as stimulating as those of the archosaurs: warm-bloodedness, extinction, descendants, transitions. Yet still no public recognition, until the past few years, when a few books, at least, have been published. Most recent in that short line is *The Evolution of Mammalian Characters*, which sets out to give an account of the origin and evolution of certain of the characters of the Mammalia, and to portray the fossils described as the living animals they once were.

The book begins with sections on vertebrate evolution from the earliest vertebrates to the cynodonts, as a brief introduction and background to the main subject. While I sympathize with the authors not wanting, for example, to get bogged down in the Devonian swamp of tetrapod origins, I found this whistle-stop survey sometimes frustratingly misleading or surprisingly archaic. For example, no mention is made of the fact that some workers propose that tetrapods might not have arisen from the crossopterygian fish, while dinocephalians are now generally considered to represent a monophyletic group. The style of writing is also somewhat dogmatic or old-fashioned in places:

"Evolution is slow, continuous and imperceptible. The evolution of major taxa does not differ in any fundamental way from the evolution of one sub-species from another" (p.49).

But the real aim of the book is to describe the origin of certain mammalian characters. Chapter 5 attempts this for dentition. At first sight this is a clear and interesting account, a potentially helpful introduction for undergraduates. Closer reading, however, reveals that there is no real attempt to address the mechanical issues involved, or the influence of muscles in the story, and the result is rather unsatisfying.

The references, too, are disappointing, especially in view of the authors' attempts to cover the literature of the past ten years. There are some surprising gaps — no reference to Bramble (1978) on jaw mechanics; hardly any to Kielan-Jaworowska's work post-1971; and no specific references to any of Kemp's papers. By contrast the illustrations are very good. Not only are many of the fossils themselves illustrated, but two particularly valuable diagrams for teaching — of standard dental terminology and cusp nomenclature — are included as well as some useful time distribution charts.

The authors say that this is a personal book, which must deflect criticism to some extent, but to me it seems that they have fallen between two stools. The book is not as comprehensive or readable as the standard undergraduate texts incorporating the mammal-like reptiles, yet it contains too many *ex cathedra* opinions to qualify as a formal review. Nevertheless, there are interesting chapters on the evolution of mammalian senses and summaries of Middle Jurassic to Upper Cretaceous mammalian faunas which students and research workers may find useful. □

Gillian King is Fellow in Zoology at St Hilda's College, University of Oxford, and Assistant Curator at the University Museum.

The atmosphere from space

J.T. Houghton

Satellite Sensing of a Cloudy Atmosphere: Observing the Third Planet.

Edited by Ann Henderson-Sellers.

Taylor & Francis: 1984. Pp. 336. \$18, \$40.

DURING the past 20 years there has been considerable growth in the exploitation of satellite platforms for mounting remote-sensing instruments to investigate the structure of the atmosphere and the Earth's surface. The great advantage of the satellite, of course, is the ease with which continuous global coverage can be achieved.

In this book, articles by different authors address various aspects of remote sensing. The first two chapters provide an introduction; radiation quantities are defined and details given of the interactions of radiation (especially visible radiation) with the atmosphere and its various components. Brief descriptions of relevant satellites are included, and there is an introductory account of cloud observations. Three of the six later contributions are concerned with various aspects of the sensing of clouds — their identification and characterization, their influence on the radiation budget, and cloud-cryosphere interactions. Other chapters deal with sounding of the atmosphere's vertical temperature profile, with problems of the sensing of land and ocean surface features (a particularly useful chapter), and with tropospheric chemistry.

To my mind, the book tends to fall

between two stools. On the one hand the chapters are not sufficiently thorough and comprehensive nor are they well enough referenced to be used as authoritative reviews. On the other hand the book lacks the cohesion and flow of exposition demanded of a textbook. There are also some surprising omissions and inclusions. Despite the introductory emphasis on the "water planet", virtually nothing is said about measurement of water-vapour and very little about precipitation measurement. Further, the importance of water in its various forms in determining the nature of the atmospheric circulation is not discussed — yet a chapter is included on tropospheric chemistry which, apart from a passing mention of a satellite instrument

(no details or results included), has no relevance to the overall satellite theme. A further annoying feature is that although in Chapter 2 a page and a half is devoted to a table giving careful definition of radiation quantities and symbols, quite different nomenclature is used in Chapter 5.

The book therefore is not, unfortunately, likely to find widespread application as a student text. But, despite its unsatisfactory features, research workers in the remote-sensing field will find that it contains some useful material and will be pleased that it is available at a price they can afford. □

J.T. Houghton is Director General of the Meteorological Office, Bracknell, Berkshire.

Unusual enzymes

Pierre Desnuelle

Lipases. Edited by Bengt Borgström and Howard L. Brockman.

Elsevier: 1984. Pp. 527. \$123, Dfl.320.

LIPASES share with phospholipases the unique ability to hydrolyse very rapidly water-insoluble substrates which form plurimolecular aggregates separated from water by an interface. Our knowledge of their properties has advanced enormously in recent years, and fully justifies the publication of a book on the topic.

Lipases was planned and partly written by B. Borgström and H.L. Brockman, long-standing experts in the field. It begins with a general account of the behaviour of lipid and water molecules at interfaces, and of the way amphiphilic lipids associate in

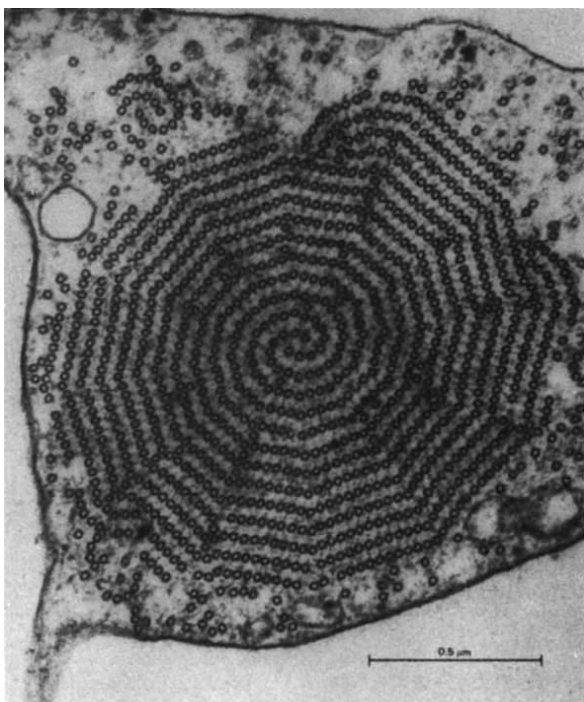
the presence of water. This useful introduction is followed by nine chapters on mammalian lipases and four which deal with "other" lipases.

Four of the early contributions are devoted to digestive lipases in mammals. The first covers lingual lipase, a newcomer to the lipase family which initiates fat splitting in the mouth and stomach, after which come two chapters on pancreatic lipase and on colipase. Pancreatic lipase has served as a model for the study of "interfacial enzymology" in heterogeneous systems; curiously enough, its interfacial adsorption is hindered by bile salts and a small protein cofactor, colipase, is required to anchor it at the interface under physiological conditions. The third digestive lipase is pancreatic carboxyl ester lipase (more commonly known as carboxylesterase or cholesterol esterase). Unlike a "true" lipase, it cleaves dissolved substrates and is activated by bile salts.

Particular emphasis in the book is laid on two tissue lipases. One is lipoprotein lipase, which acts at the capillary endothelium to remove circulating chylomicrons and VLDL from the plasma, and is activated by apolipoprotein CII. This enzyme is also present in milk, together with an enzyme similar to carboxylesterase. The other tissue lipase covered in some detail is adipose tissue lipase (also called hormone-sensitive lipase) which is involved in the mobilization of triglycerides from the adipose tissue, its action being controlled by a cyclic AMP-mediated phosphorylation. Finally, there are accounts of lipases from plants, fungi and bacteria, special reference being made to structure, specificity and practical applications.

Most of the chapters are authoritative yet readable, and each of them is followed by a thorough list of references. As well as its natural audience of enzymologists and biologists engaged in research on lipid metabolism, the book will interest clinicians and advanced students in biology and medicine. □

Pierre Desnuelle is Professor Emeritus of Biochemistry at the University of Marseilles.



Sight for the cell tourist — electron micrograph of a section through an axopod of the heliozoan *Echinospheerium nucleofilum*, showing the double-helical arrangement of microtubules with twelve-fold symmetry. The illustration is reproduced from *A Guided Tour of the Living Cell* by Christian de Duve, to be reviewed in *Nature* next month.

New products listed this week include a simple magnetic susceptibility balance and new software for the Apple II microcomputer.

● The new Hybritech Photon is a combined computerized spectrophotometer-spectrofluorometer designed to perform immunoassays. All readings are performed bichromatically, thereby allowing measurement in the assay tube and thus reducing time and improving accuracy. Three primary wavelengths are available and these are automatically programmable through the keyboard. The microprocessor has a 64K memory and is operated through an alpha-numeric keyboard.

Circle No. 100 on Reader Service Card.



Crucibles in molybdenum and titanium from B-J Scientific

● B-J Scientific Products of Albany, Oregon, has extended its range of laboratory crucibles to include molybdenum and tantalum vessels for use with strong mineral and organic acids and salts, and in the handling of molten metals. Tantalum displays corrosion resistance to molten metals such as magnesium, potassium, lithium, and sodium at up to 1,100°C and is highly resistant to most acids. Molybdenum crucibles resist many acids and molten metals. The crucibles are available in sizes from 5ml to 1,000ml, with or without covers.

Circle No. 101 on Reader Service Card.

● Labpack data logging package enables the Apple II microcomputer to function as a laboratory data logger and a processor at relatively low cost. The Labpack includes hardware for converting analog signals into digital form for storage in the computer, and software to operate and control the converter and to provide comprehensive procedures for data processing and graphical presentation. Labpack is available from Adept Scientific Micro Systems of Hitchin, Herts.

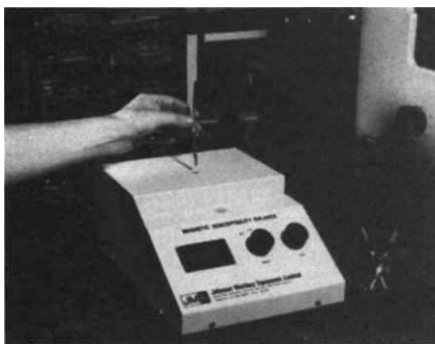
Circle No. 102 on Reader Service Card.

● The IMPACT system for automated non-isotopic immunoassay is now available from Acade SA, Brussels, Belgium, with eight test kits — for T4, T3, TSH, TBG, ferritin, alpha-fetoprotein, human placental lactogen and C-reactive protein. IMPACT is an acronym for immunoassay by particle counting, a non-isotopic technique which brings the level of substances which can be determined automatically in the clinical laboratory down to 10^{-15} molar.

Circle No. 103 on Reader Service Card.

● The FRAC-100 fraction collector from Pharmacia Fine Chemicals now features a new solvent resistant bowl, the FRAC-100 Bowl SR. This will be particularly useful in its application to HPLC. Constructed from Valox R, the bowl is resistant to most common organic solvents, strong acids, bases and aqueous buffers of any pH.

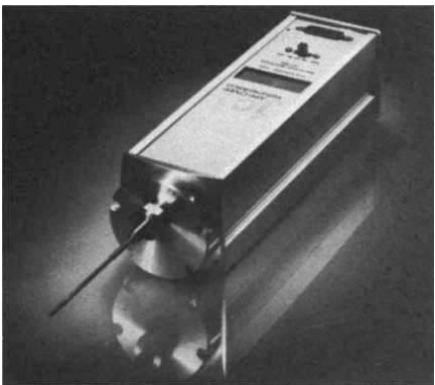
Circle No. 104 on Reader Service Card.



Simple testing for magnetic susceptibility — Johnson Matthey's new balance

● Johnson Matthey Equipment is now manufacturing a new magnetic susceptibility balance for use as a research tool and as a teaching aid. Unlike conventional systems, the JME balance makes use of magnets and a stationary sample tube. Magnetic susceptibility is calculated from an instantaneous digital readout by using a simple equation. Sample weights of around 250 mg are all that are required to give accurate measurements.

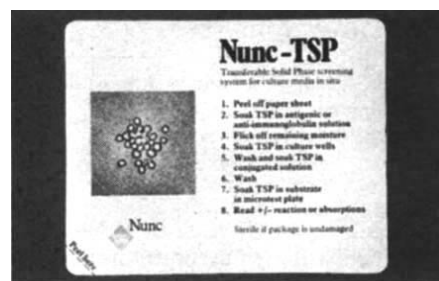
Circle No. 105 on Reader Service Card.



The Chubb microprobe for measuring static

● Two new devices from John Chubb Instruments — the JCI 201 and 202 Electrostatic Mini-Probes — make it possible for the distributions of static potential and charge to be studied in and around small scale structures inaccessible to conventional instruments. The probes give stable and accurate measurement of local potential in the vicinity of static charges on insulating and conducting surfaces. Both probes can resolve potentials to 1 volt and the 201 probe may be used fully immersed within insulating liquids.

Circle No. 106 on Reader Service Card.



New from Nunc for monoclonal screening

● A new transferable solid phase screening system from Vanguard/Nunc is designed to simplify the screening of monoclonal antibodies in hybridoma culture. The Nunc-TSP Screening System uses the ELISA method but uses a specially constructed TSP plate with 96 pins that fit into a 96-well microtestplate. The TSP plate is treated for optimum protein adsorption to the exterior walls of the pins. The TSP plate is placed directly into the microtestplate, so the testing procedure does not affect cell growth, without changing the composition of the media. A Nunc TSP Wash Unit is available for rapid cleaning of the pins. It comes in a semi-automatic or a fully automated version with an electronic control unit.

Circle No. 107 on Reader Service Card.



Falcon tubes in their disposable polystyrene rack

● The Falcon Labware Division of Becton Dickinson has introduced improved conical tubes with printed graduations formulated from a new form of clear polystyrene. The new material has high contact transparency and see-through clarity providing a clear view of the specimen, fluid level and graduation markings. The tubes have been tested under normal usage conditions at 9,400g, and the graduations are in blue for easy volumetric identification. The new material's stability allows radiation sterilizing. The tubes are available in 15-ml and 25-ml sizes and can be obtained in disposable polystyrene rack with convenient tube-spacing and window slots to permit a clear view of each tube.

Circle No. 108 on Reader Service Card.

These notes are based on information provided by the manufacturers. For further details circle the appropriate numbers on the Reader Service Card bound inside the journal.